75th MORSS CD Cover Slide

712CD

For office use only 41205

UNCLASSIFIED DISCLOSURE FORM CD Presentation

12-14 June 2007, at US Naval Academy, Annapolis, MD

Please complete this form 712CD as your cover page to your electronic briefing submission to the MORSS CD. Do not fax to the MORS office.

Author's Request: I request authority to disclose the following presentation in the MORSS Final Report, for inclusion on the MORSS CD and/or posting on the MORS Website.

Name of Principal Author and all other author(s): Gregory T. Hutto, YD-3 DAF

Principal Author's Organization and address:

Phone:_850,882,0607_____________

Fax:____850.882.5644____________

Email:___gregory.hutto@eglin.af.mil____

Original title on 712 A/B:___________Design of Experiments_____________________________

Revised title:____

Presented in: ***Tutorial Session***

This presentation is believed to be:
Distribution A: UNCLASSIFIED AND APPROVED FOR PUBLIC RELEASE

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 JUN 2007**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Elements of Design of Experiments**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **53d Test Mgt Group Eglin AFB; Florida State University, Sverdrup**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

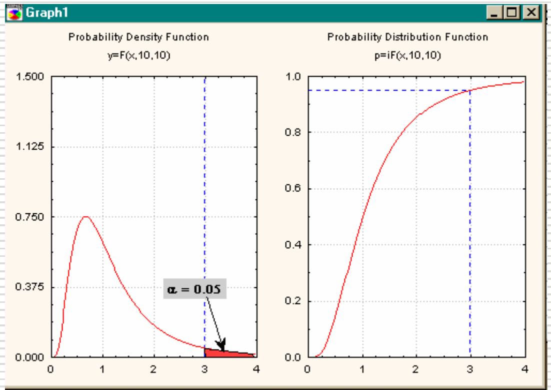
13. SUPPLEMENTARY NOTES
See also ADM202526. Military Operations Research Society Symposium (75th) Held in Annapolis, Maryland on June 12-14, 2007, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**146**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

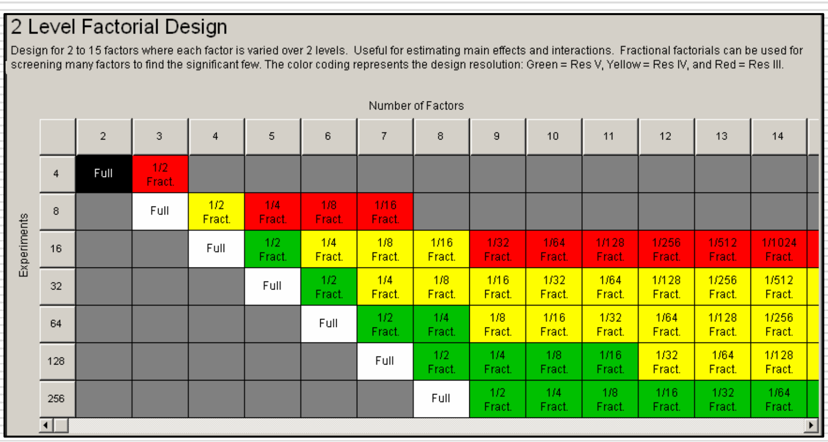


Elements of Design of Experiments

Greg Hutto, 53d Test Mgt Group

Jim Simpson, Florida State University, Sverdrup

- Introductions
- Course Overview
- Exercise



2 Level Factorial Design

Design for 2 to 15 factors where each factor is varied over 2 levels. Useful for estimating main effects and interactions. Fractional factorials can be used for screening many factors to find the significant few. The color coding represents the design resolution: Green = Res V, Yellow = Res IV, and Red = Res III.

Number of Factors (columns) / Experiments (rows)

	2	3	4	5	6	7	8	9	10	11	12	13	14
4	Full	1/2 Fract.											
8		Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.							
16			Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.	1/512 Fract.	1/1024 Fract.
32				Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.	1/512 Fract.
64					Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.
128						Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.
256							Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.

Course Overview

- Session 1 – Central challenge of test
  - Deep and broad testing
  - What is DOE?
  - History
- Session 2 – Comparing test strategies, which one is best?
  - Conventional logic
  - Designed experiment
- Session 3 – Factorial designs
  - Phases of conducting DOE tests
  - Project decomposition
  - Planning the test
  - Producing observations
  - Pondering results
- Session 4 – Fractional factorials and other topics

Course Objectives

- Objectives:
  - Understand Test Stats through ANOVA
    - Plots and Displays
    - Inferences - Sampling Distributions
    - Multi-way experiments
    - Fractional Replication
    - ANOVA Extensions and Modifications
  - Understand classical model and assumptions
    - Create and analyze two-level experiments
    - Detect and correct violations of assumptions
    - Estimate sample size and formulate designs
  - Be able to use Excel, Resampling and Design Ease for Basic Design and Analysis

Background -- Greg Hutto

- B.S. US Naval Academy, Operations Analysis
- M.S. Stanford University, Operations Research
- USAF -- TAWC Green Flag, AFOTEC Lead Analyst
- Consultant -- Booz Allen & Hamilton, Sverdrup Technology
- Technical Fellow in Operations Research -- Sverdrup Technology
- TMG OA and DOE Champion – 53rd Wing
- Design of Experiments -- 15 Years

Selected T&E Project Experience -- 15+ Years

- Green Flag Ex '79
- F-16C IOT&E '83
- AMRAAM, JTIDS ' 84
- NEXRAD, CSOC, Enforcer '85
- Peacekeeper '86
- B-1B, SRAM, '87
- MILSTAR '88
- MSOW, CCM '89
- Joint CCD T&E '90
- SCUD Hunting '91

- AGM-65 IIR '93
- MK-82 Ballistics '94
- Contact Lens Mfr '95
- 30mm Ammo '97
- 60 ESM/ECM projects '98--'00
- 200 projects B-1B SA OUE, Maverick IR+, F-15E Suite 4E+,Chem Detector '01-'05

Background – Jim Simpson

- B.S. US Air Force Academy, Operations Research
- M.S. Air Force Institute of Technology, Operations Research
- Ph.D. Arizona State University, Industrial Engineering
- USAF – Armament Division OA, Personnel Analyst, Faculty USAFA
- Consultant -- Sverdrup Technology
- Florida A&M – Florida State University, Industrial and Manufacturing Engineering

Project Experience

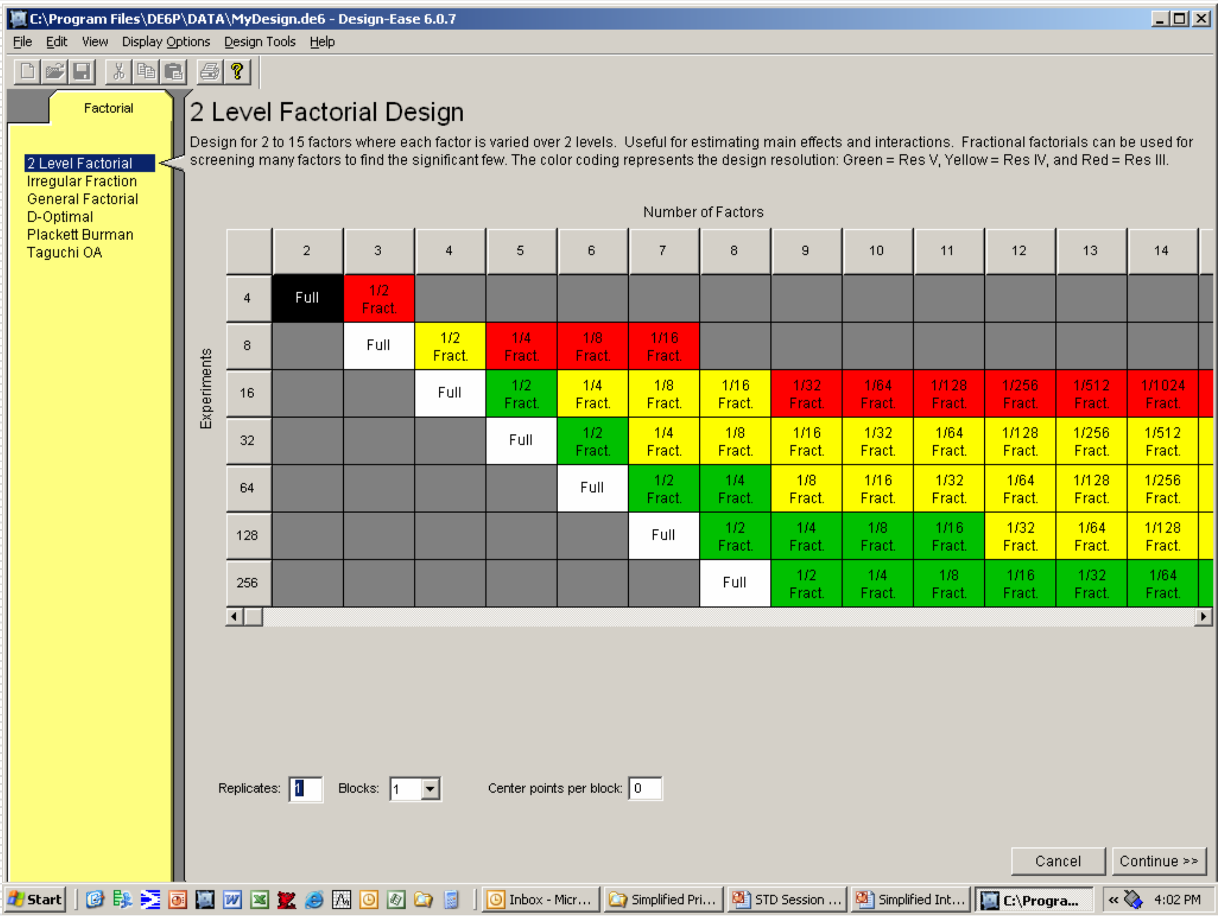
- CV-22
- AGM Simulation
- WMD Effectiveness
- Army BDA
- Army Operator Effectiveness

- 53rd TMG Consultant
- Wind Tunnel Testing
  - Air Force
  - NASA
  - NASCAR

Class Introductions

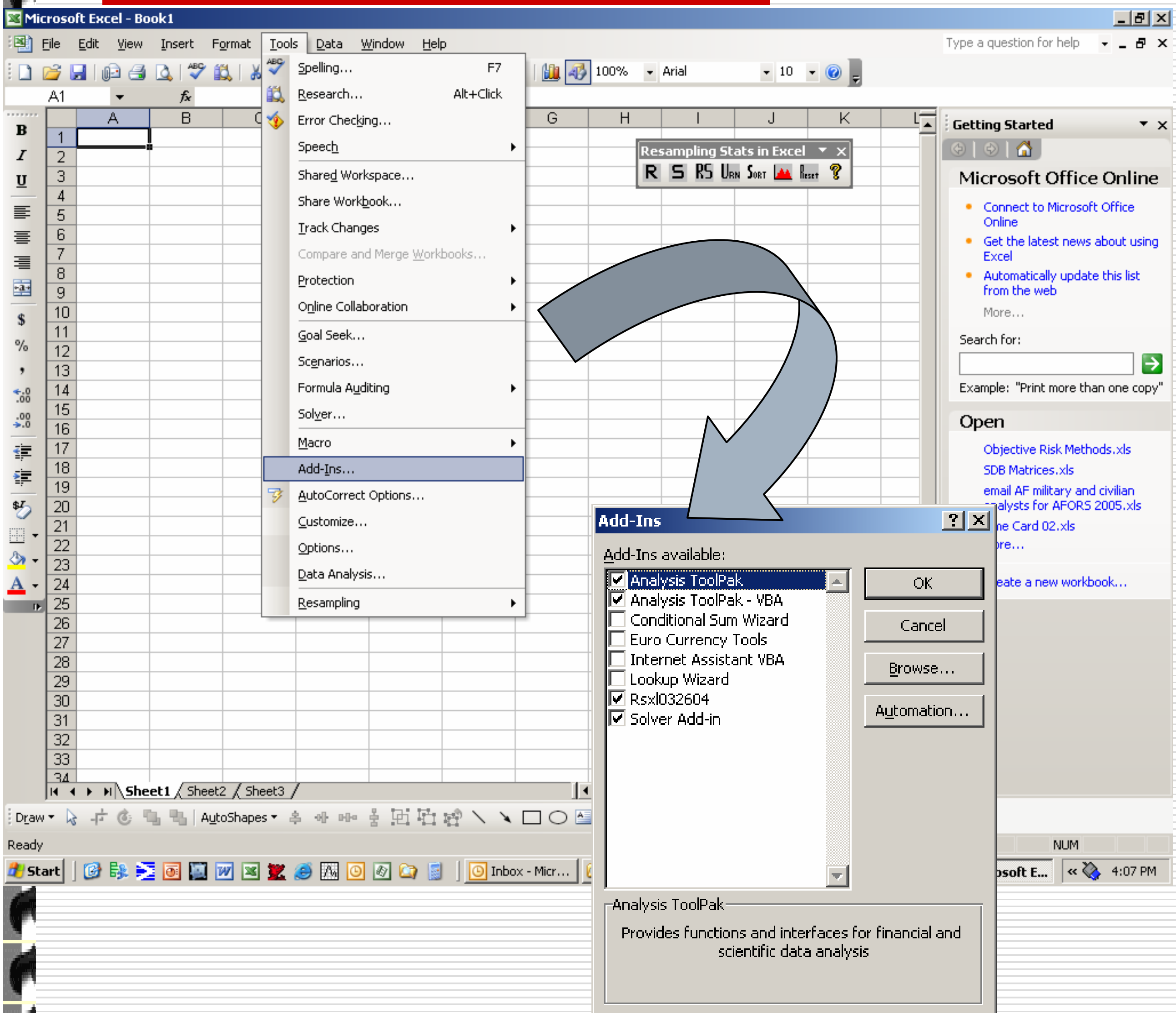
- Name
- Degree(s) + Field of Study
- Test experience –
  - types and years
- Had any Stats?
  - If *yes*, ever used it?

Loading Design Ease



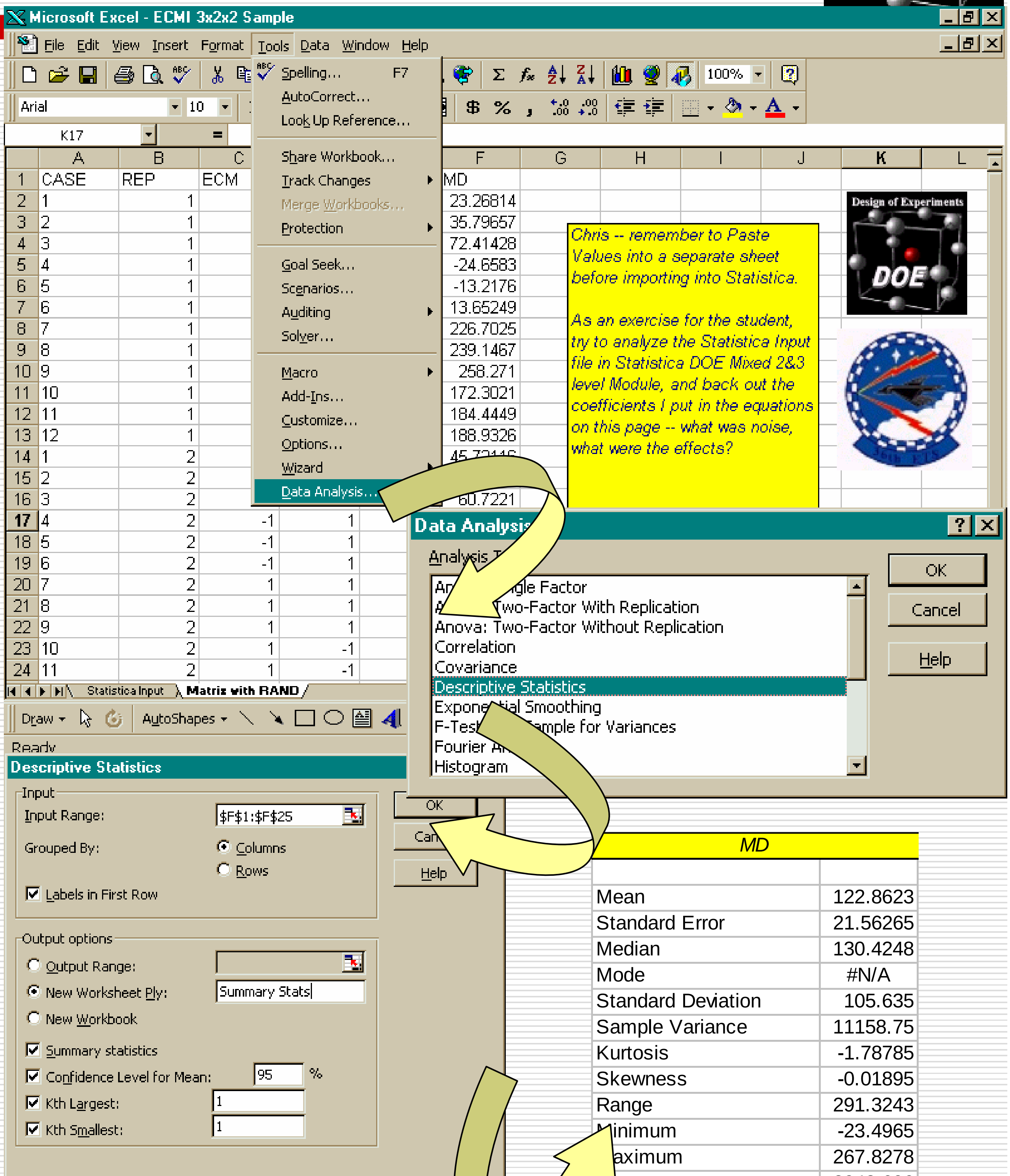
- Standard Auto Run Install
- 30 day limit student version
- Powerful, yet easy-to-use DOE package

Loading Excel's Analysis ToolPack



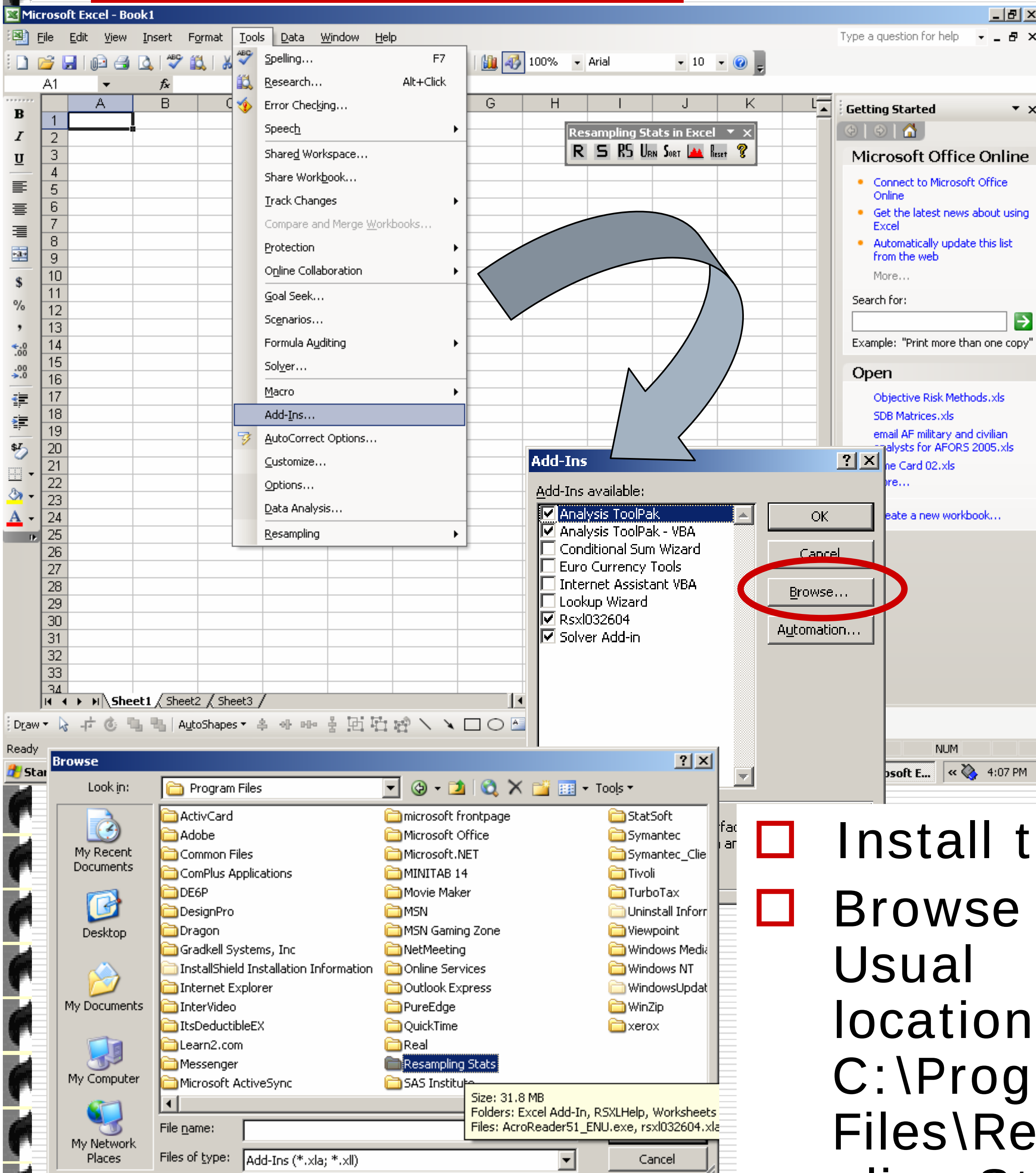
- Check first two checkboxes-
  - Analysis ToolPack + Analysis ToolPack VBA
  - Required for Basic Stats and Resampling Stats

Loading Excel's Analysis ToolPak



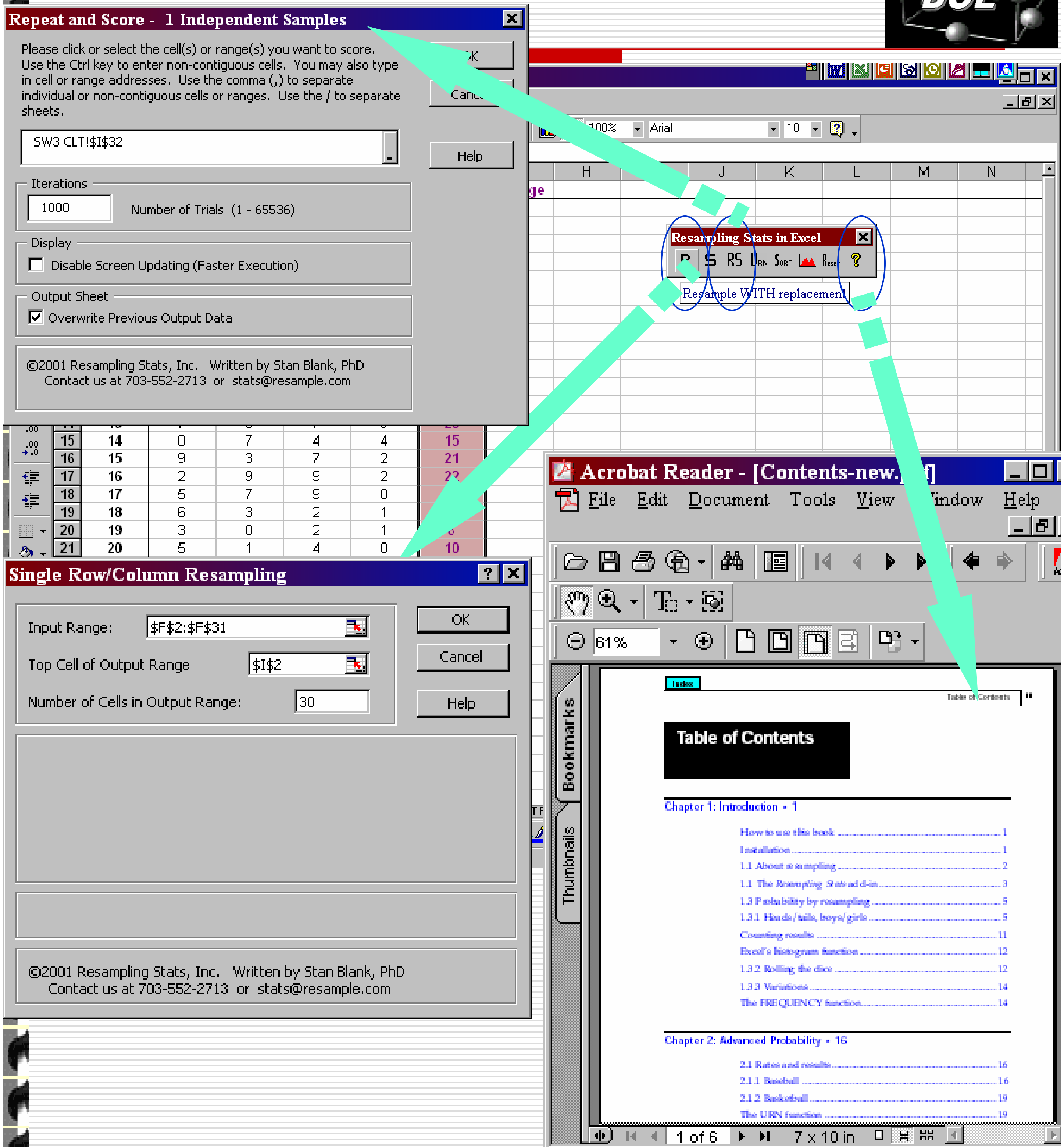
MD	
Mean	122.8623
Standard Error	21.56265
Median	130.4248
Mode	#N/A
Standard Deviation	105.635
Sample Variance	11158.75
Kurtosis	-1.78785
Skewness	-0.01895
Range	291.3243
Minimum	-23.4965
Maximum	267.8278
Sum	2948.696
Count	24
Largest(1)	267.8278
Smallest(1)	-23.4965
Confidence Level(95.0%)	44.60568

Loading Resampling Stats



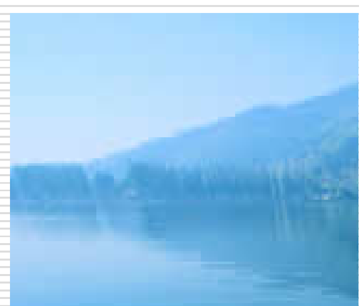
- Install then
- Browse … Usual location C:\Program Files\Resampling Stats

Resampling Stats – ?, R and RS



News Flash! US Company Indicted for Iraq Arms Sales!

- The SAM-in-a-Box is a potent threat!
- Sold to Saddam by Stat-a-Pult Inc.
- We must conduct an exploitation HWIL simulation today
- MOP – miss distance
- Conditions – projectile, propellant, Target RCS & velocity, ECM … etc
- To start … Stopper Setting @ 2 levels



Planning a Test

Objective: Develop an experimental design for understanding the effect of stop position on the distance of a projectile travels when launched from a catapult, hereafter called a 'statapult.'

Background: Although the statapult has several factors that may affect distance, we will only consider stop position, and specifically just two settings for stop. Your task will be to design a test that will capture the influence of changing stop position on distance traveled.

Direction: Develop a test plan for determining how stop position affects projectile launch distance. All other statapult settings will be specified as fixed. Your mission is to decide the number of tests required, the order in which to run the tests, and how you plan to assess the data after test. You have 10 minutes.

# of Runs	Run Order	Analysis Approach

Example or Partial Test Matrix

Run	Stop Setting	Distance
1		
2		
3		
4		
5		
6		
7		
. . .		
n		

Single Factor Design

Conduct your test to determine the effect of stop position on distance. The first four planning steps are completed below.

Objective	Determine the effect of stop position on distance.
Factor(s)	The input factor is stop position. Stop position will have two settings.
Response	The output is distance in inches. Use a launch angle of 172 egrees.
Design	The following table is the design matrix. Perform the experiments below.

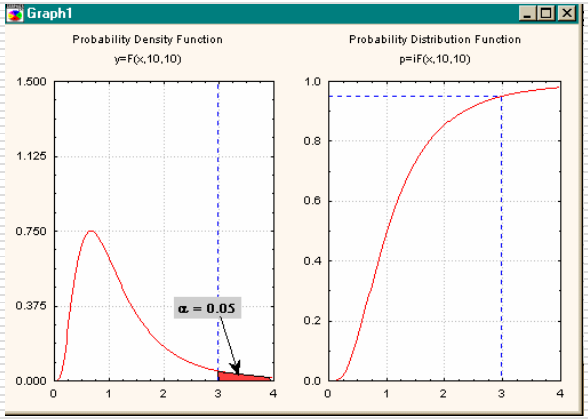
Run	Stop Position	Distance
1	3	
2	2	
3	2	
4	3	
5	2	
6	3	
7	2	
8	3	

Assignment:

1. Each team will have a launcher, a holder, a recorder and range finders. Teams are not permitted to communicate with other teams.

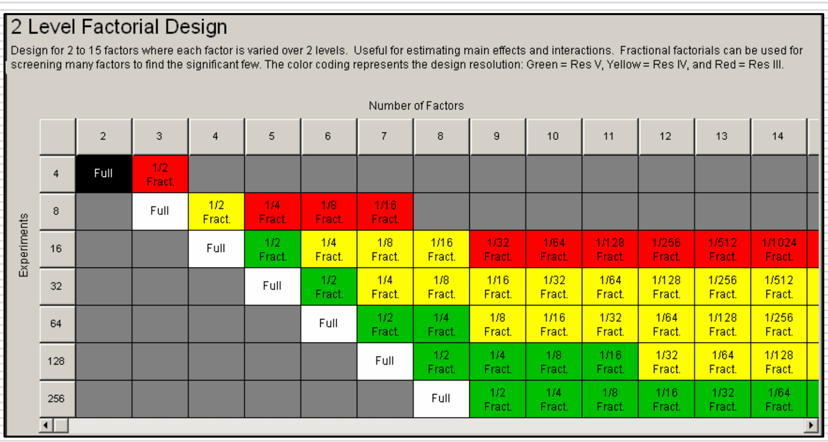
2. Collect data according to the above design matrix. Each team will have **3 minutes to organize themselves and collect the data**.

3. We will step you through an analysis of your data.



Session 1 Testing Deeply and Broadly

- The General Test Problem
- Testing Deeply – 1 Sample
- Review of Descriptive Stats
  - Basic numerical statistics
  - Pictures of the data
- Testing Broadly - Spiral 1
  - Design for Maverick H/K



2 Level Factorial Design

Design for 2 to 15 factors where each factor is varied over 2 levels. Useful for estimating main effects and interactions. Fractional factorials can be used for screening many factors to find the significant few. The color coding represents the design resolution: Green = Res V, Yellow = Res IV, and Red = Res III.

Number of Factors

Experiments	2	3	4	5	6	7	8	9	10	11	12	13	14
4	Full	1/2 Fract.											
8		Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.							
16			Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.	1/512 Fract.	1/1024 Fract.
32				Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.	1/512 Fract.
64					Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.	1/256 Fract.
128						Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.	1/128 Fract.
256							Full	1/2 Fract.	1/4 Fract.	1/8 Fract.	1/16 Fract.	1/32 Fract.	1/64 Fract.

A beer and a blemish …

- 1906 – W.T. Gossett, a Guinness chemist
- Draw a yeast culture sample
- Yeast in this culture?
- Guess too little – incomplete fermentation; too much -- bitter beer
- He wanted to get it *right*

- 1998 – Mike Kelly, an engineer at contact lens company
- Draw sample from 15K lot
- How many defective lenses?
- Guess too little – mad customers; too much -- destroy good product
- He wanted to get it *right*

The central test challenge …

- In all our testing – we reach into the bowl (reality) and draw a sample of operational performance
- Consider our "Sam-in-a-Box"
  - Suppose an historical 70% hit rate
  - Is this captured version at least as good?
- We don't know in advance which bowl God hands us …
  - The one where the system *works* <u>or</u>,
  - The one where the system *doesn't*

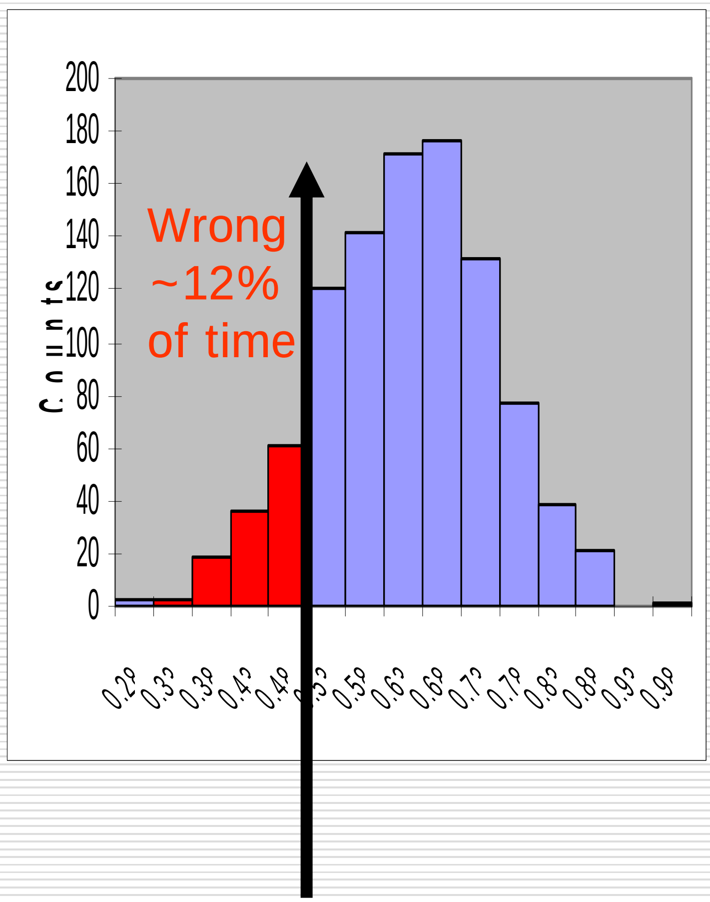
Start -- Blank Sheet of Paper

- Let's draw a sample of _*n*_ shots
- How many is enough to get it *right*?
  - 3 – because that's how much $/time we have
  - 8 – because I'm an 8-guy
  - 10 – because I'm challenged by fractions
  - 30 – because something good happens at 30!
- Let's start with 20 and see ...

A False Positive – Declaring StataPult is Degraded (when it's not) -- α

- Suppose we *fail* StataPult when we get 50% hits
- We'll be wrong (on average) about 12% of the time
- We can tighten the criteria (fail on 60%) by failing to field more good weapons
- We can loosen the criteria (fail on 40%) by missing real degradations
- Let's see how often we miss such degradations ...

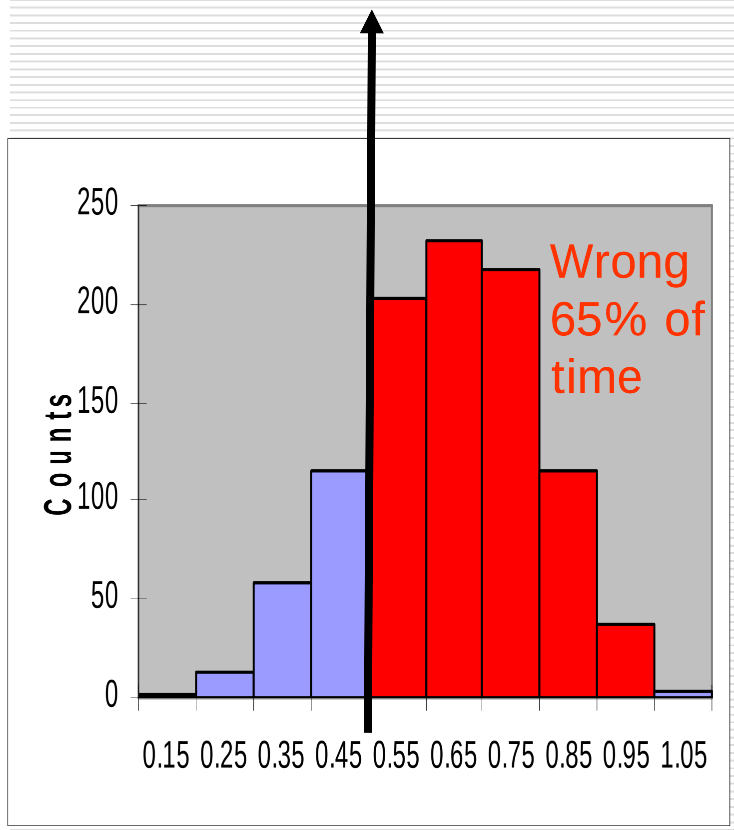
In this bowl – Statapult performance **meets** spec



A False Negative – Declaring Statapult Meets Spec (when it doesn't) -- β

- Use the failure criteria from the previous slide
- If we *field* Statapult with 55% hits, we fail to detect the true degradation
- If Statapult fails spec, with n=20 shots, we'll miss it about 80% of the time
- We can, again, tighten or loosen our criteria, but at the cost of increasing the other error

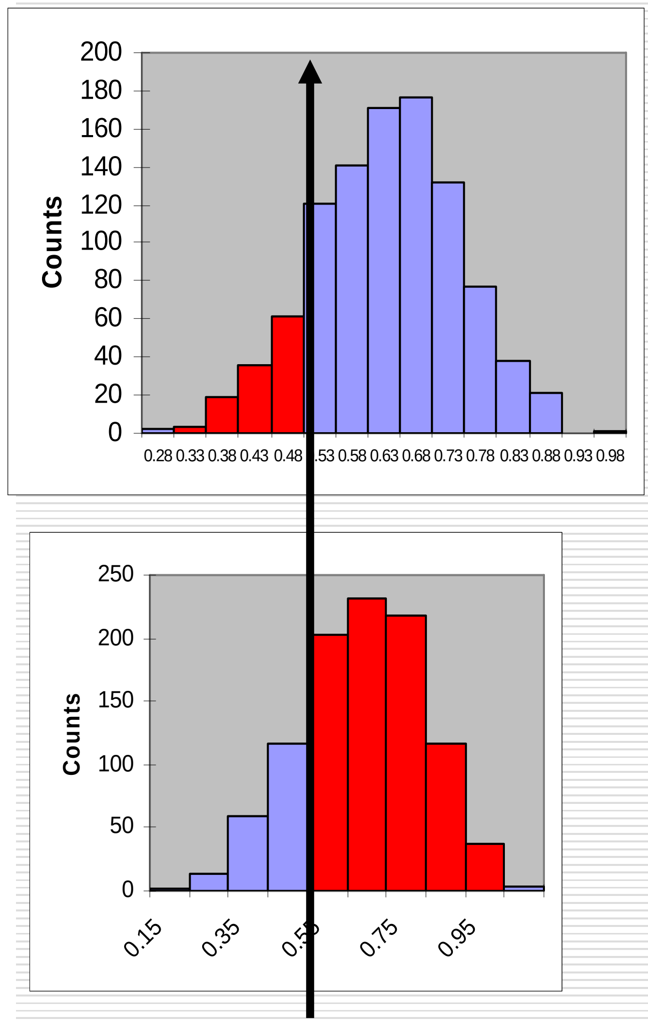
In this bowl – Statapult Prob(in) **decreased** 10% from 70% to 60%- it is degraded



Q1 – How Many Runs?
Testing Deeply Enough to Balance Our Chance of Errors

Design of Experiments
DOE

- Putting them together, we see we can trade one error for the other (α for β)
- We can also increase sample size to decrease our risks in testing
- These statements are not opinion – they are <u>mathematical fact</u> and an inescapable challenge in testing
- There are two *other* ways out … factorial designs and real-valued MOPs

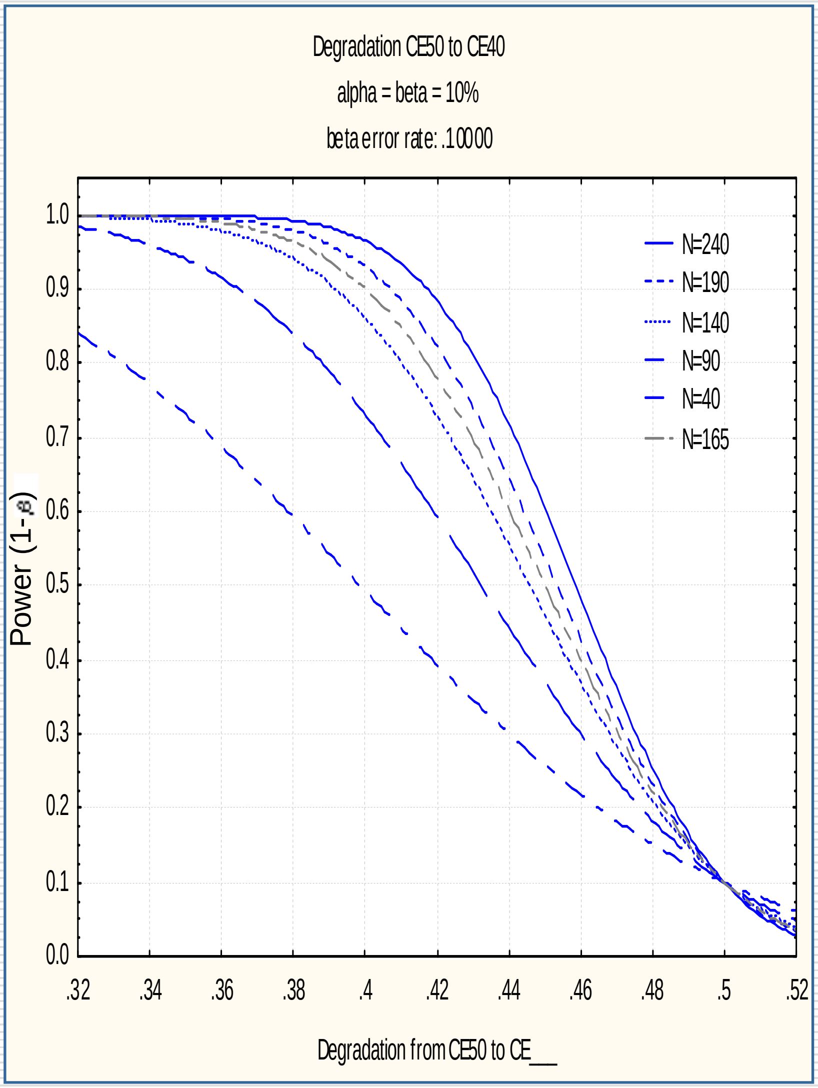


Getting it right: ***Confidence*** in stating results; ***Power*** to find small differences

Q1 – How Many Runs?

- We teach our analysts and test engineers to custom-design sample size for each test
- The answer is – it depends on:
  - Appetite for risk
  - Difference we care about
  - Amount of noise in the system under test

The General Answer -- Power curve for this test



Recap – First Two Challenges

- Challenge 1: effect of sample size on errors – Depth of Test
- Challenge 1a: measuring best response – rich and Relevant
- So -- it matters how many we do and it matters what we measure
- There is a 2nd challenge – Breadth of testing – searching the employment battlespace

SAM-in-a-Box HWIL .. continued

Step 2. Now add delivery method (arm draw-back distance) and target RCS (stealth or conventional) to the test conditions. Revise your test matrix and reconsider the previous questions (trials, order, analysis)

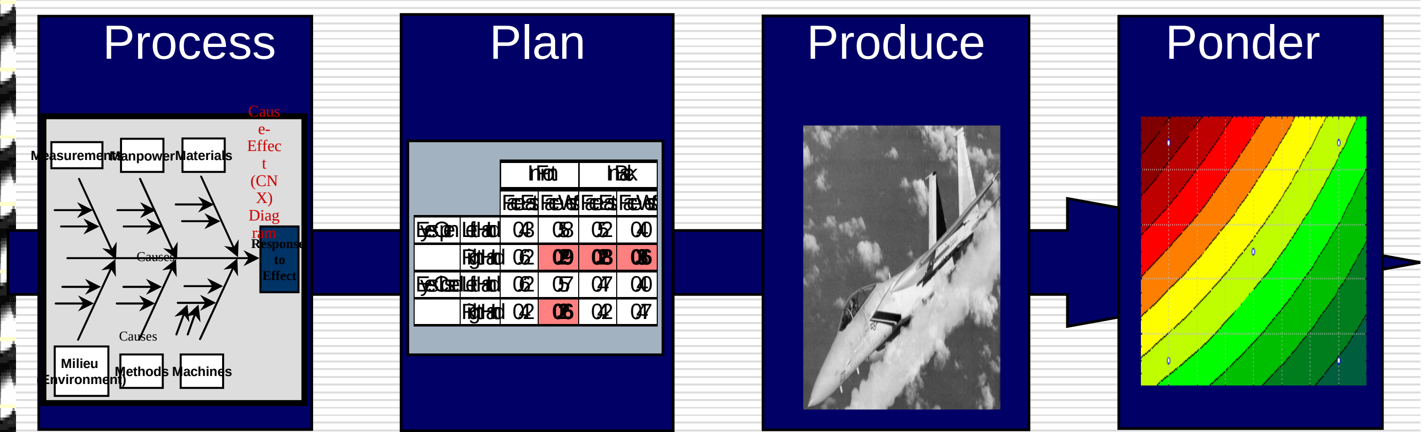
You might be interested in answering questions like:

1. What conditions produce the best accuracy on the target?
2. What conditions are robust to uncontrollable factors (jamming, fatigue, chaff, haze)?
3. What conditions fail to make a difference in accuracy?

With these three test conditions, how will you revise your test to answer the above questions?

This situation is exactly that faced by our test teams. In this course, we will learn about a method to answer them…

Spiral 1 --How We Design Tests in 4 Stages



I Project Description and Decomposition

II Plan the Test Matrix

III Produce the Observations

IV Ponder the Results

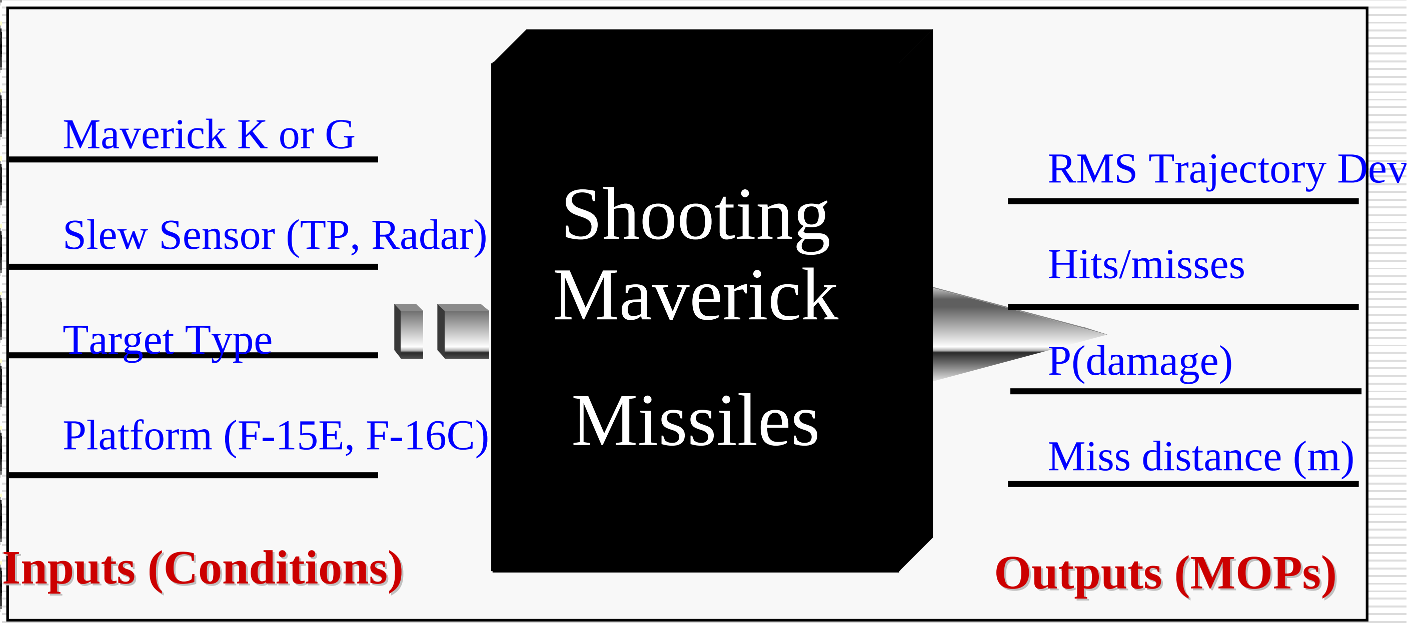
We'll stroll quickly through examples of each step for Maverick Problem

Switch Examples: The Maverick H/K

- Present B (EO) and D/G (IIR) seekers are degrading with age
- Replace EO/IR versions with new H and K variants
- Critical Operational Issue: Does Maverick still perform at least as well as previously?
- For our purposes – we shoot several Mavericks of each type – same day, same conditions, same target
- What are our chances of getting it *right*?

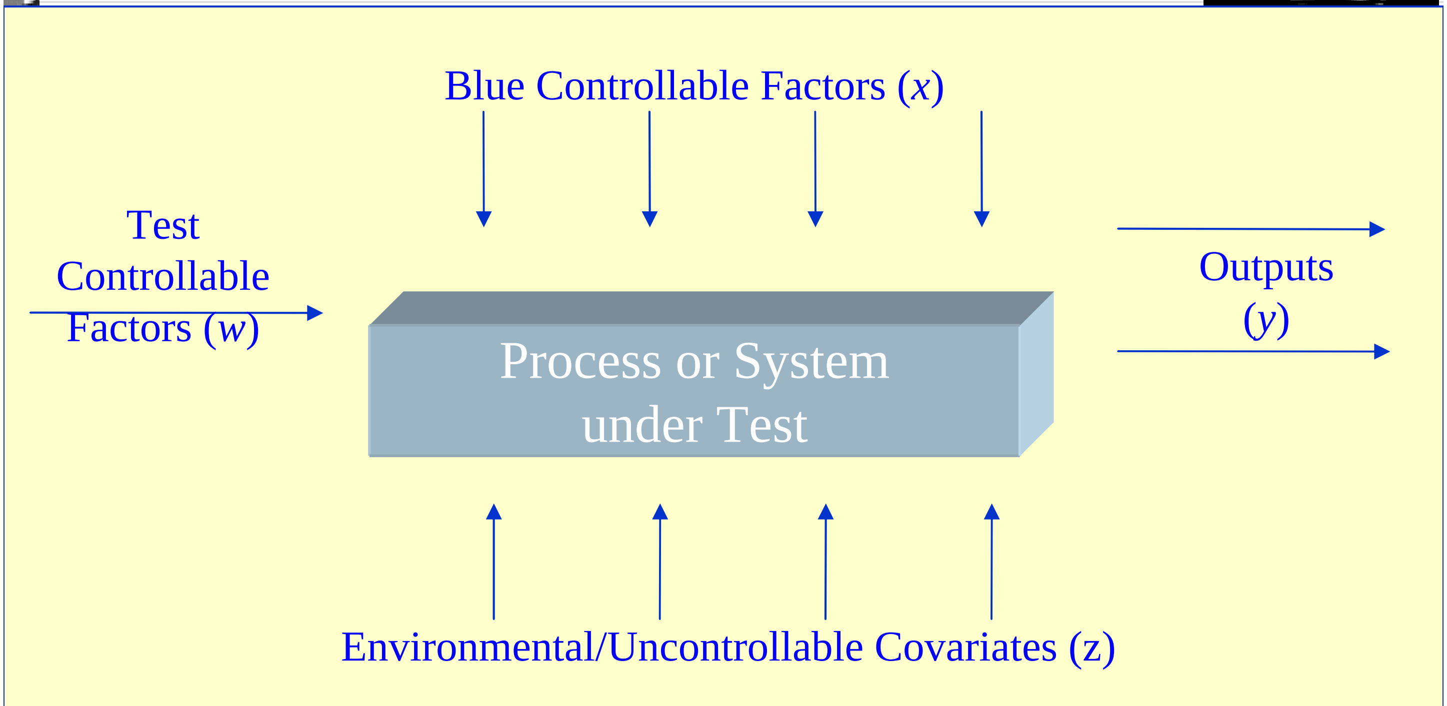
Challenge 2: Broader -- How Do Designed Experiments Solve This?

Designed Experiment *(n). Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).*



G.E.P Box (son-in-law to Fisher) –
"All math models are false ...
but some are useful."
"All experiments are designed ...
most, poorly."

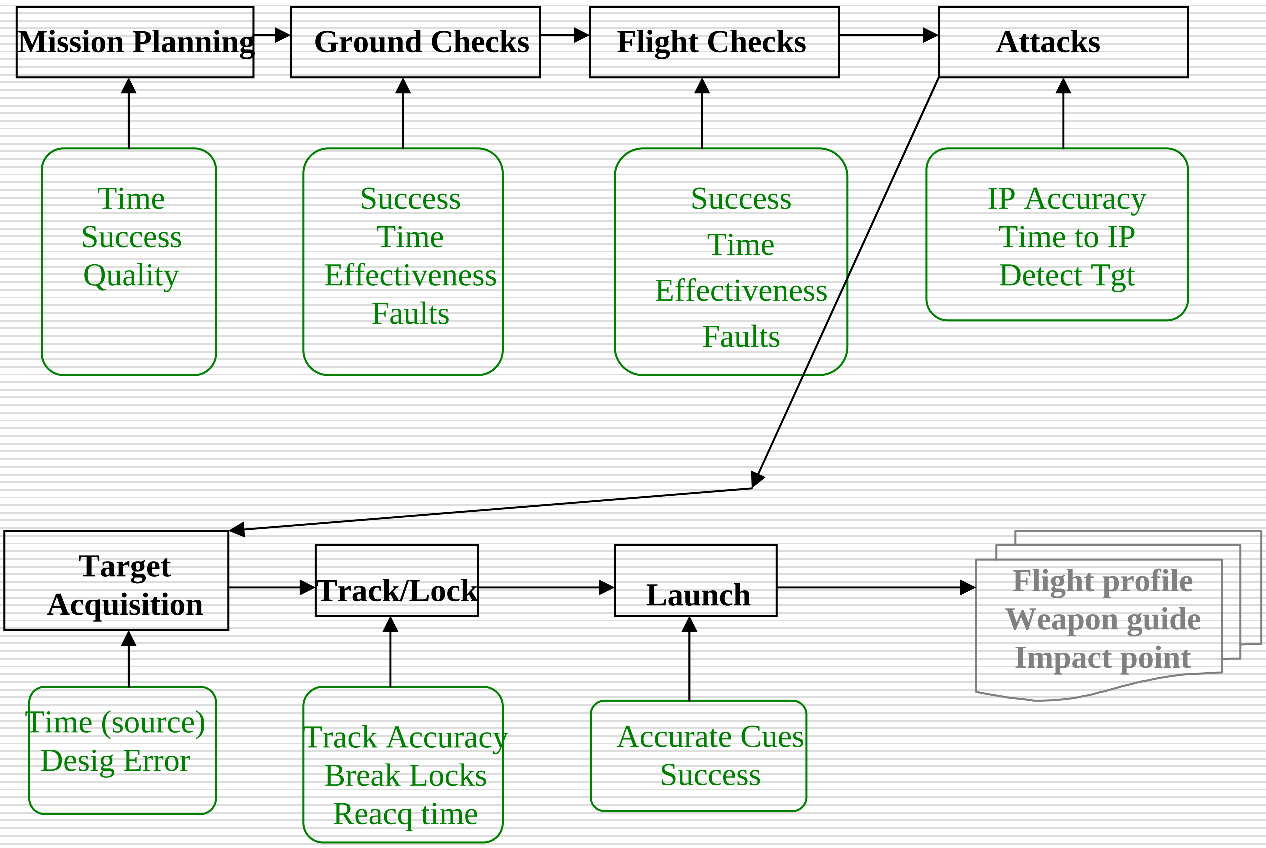
The General Test Problem – A Little More Complicated Black Box



- Questions that must be answered for any test:
  - Which of the variables (x's, w's) to change, what pattern, what range? How deal with z's?
  - How many runs should we/can we make?
  - How should the runs be sequenced?
  - What method of analysis should be employed?
- Possible Objectives:
  - Characterizing where y's usually fall depending on x settings
  - Determining which variables have or do not have an influence on outputs
  - Determining where to set influential variables so outputs are near goal
  - Determining where to set x's so that effects of z's are small

Answers determine your **strategy of experimentation**.

PROCESS Process Flow: Steps in Black Box – MOPs to measure



- The test team (Ops, Analyst, PM, eng) breaks process down
  - What are the events, steps, **outcomes**?
  - What are the ops choices and conditions at each step?
  - How do we measure success?
- Result – Results to Measure

Cause and Effect Diagram to Elicit Maverick Shot Test Conditions

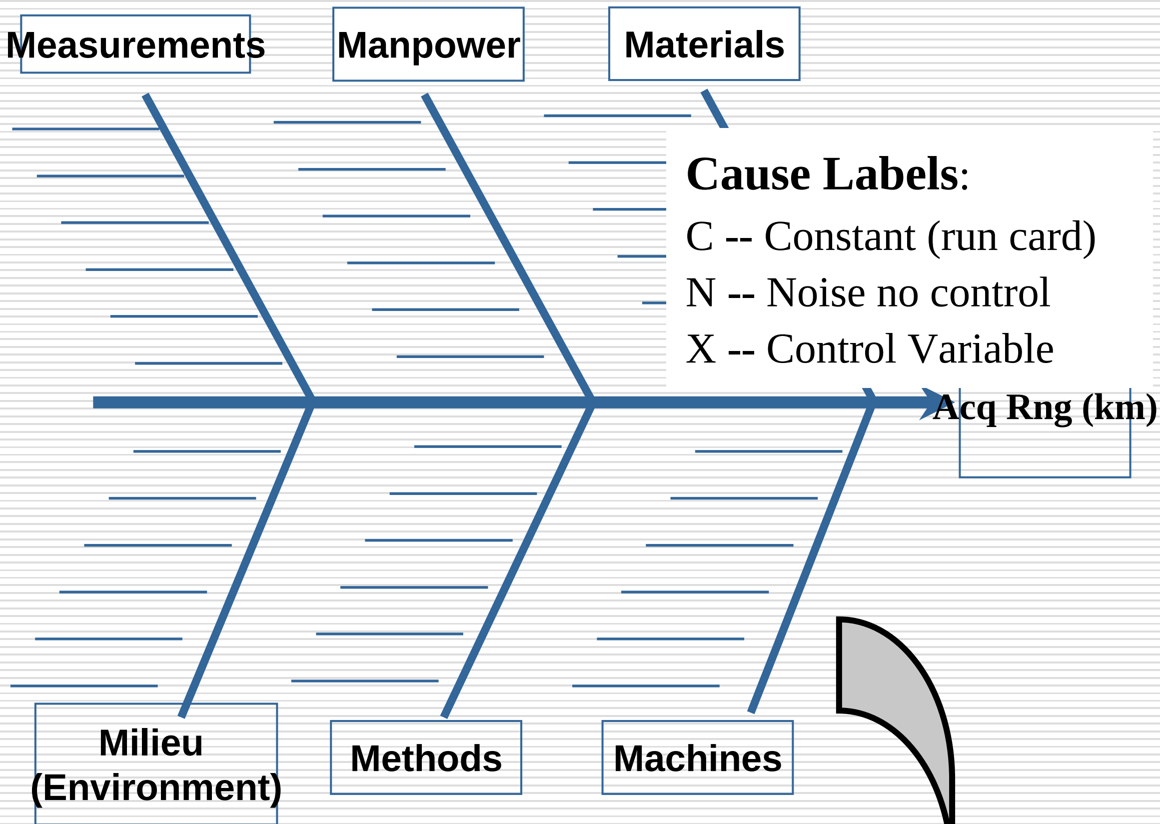


Table of Conditions

N	Variable	Units	Range	Priority	Exp. Control	Design Range
1	Target	Qual	many	H	X	trk, tank, bld
2	Platform	Qual	F-15/16	H	X	F-15, F-16
3	Time of Day	hour	0-24	H	X	Dawn, Noon
4	Humidity	gm/m3	30-Jun	H	X	Nellis Eglin
5	Tgt Velocity	mph	0 - 40	H	C	stationary
6	Operator Skill	Qual	L,M,H	M	C	H- captive runs
7	Mav Model	Qual	BDGHK	L	C	B-H or D-K
8	Wind	m/s	0-10	L	N	measure, R

Battlespace Measures & Conditions for Maverick H/K Case

Design of Experiments
DOE

Measures of Performance

Type	Measure of Performance
Objective	Target acquisition range
	seeker lock-on range
	launch range
	mean radial impact distance
	probability of target impact
	reliability
Subjective	Interoperability
	human factors
	tech data
	support equipment
	tactics

Test Conditions

Conditions	Settings	# Levels
Missile Variant:	H, K, B, G	4
Launch Platform:	F-16C, F-15E, A-10A	3
Launch Rail:	LAU-117, LAU-88	2
Target:	Point, Area	2
Time of Day:	Dawn/Dusk, Mid-Day	3
Environment:	Forest, Desert, Snow	3
Weather:	Clear (+7nm), Haze (3-7nm), Low Ceiling/Visibility (<3000/3nm)	3
Humidity:	Low (<30%), Medium (31-79%), High (>80%)	3
Attack Azimuth:	Sun at back, Sun at beam, Sun on nose	3
Attack Altitude:	Low (<5000'), High (>5000')	2
Attack Airspeed:	Low (Mach .5), Medium (Mach .72)	3
Missile Mode:	Centroid track, Force correlate track	2
	Combinations	139,968

Obviously a large test envelope … how to search it?

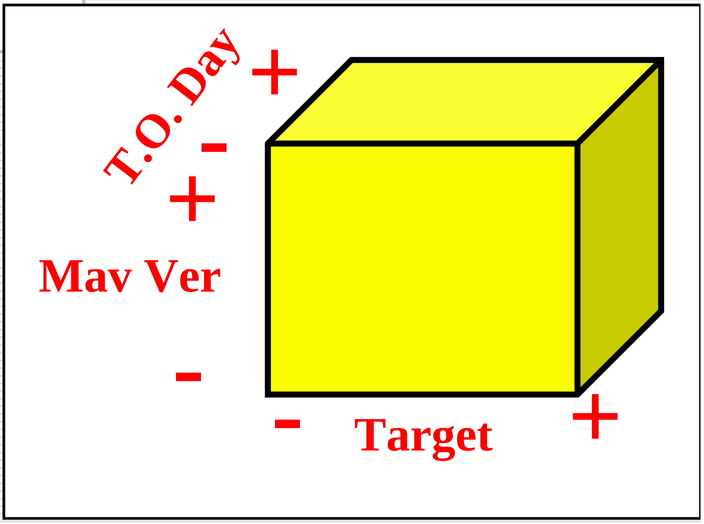
PLAN: Algorithm to Construct Full Factorial Run Matrices

Case	TimeofDay	Maverick	Target	Aircraft
1	0500	H	Building	F-15E
2	0500	H	Building	F-16C
3	0500	H	Tank	
4	0500	H	Tank	
5	0500	K		
6	0500	K		
7	0500	K		
8	0500	K		
9	1200			
10	1200			
000	000			
23	1830			
24	1830			

- How many Variables (A,B,C … K)?
- How many settings (levels) for each Variable (a, b, c, … k)?
- How many runs total?
  - a *b* … *k
  - Example: 4 vars: 1 at 3, 3 at 2 levels
  - Runs (N) = 3*2*2*2 = 3 * 2^3 = 24 runs (test cells)
- Construct Design --
  - Var A set at 1 for N/a, level 2, etc.
  - With A at 1, set B at 1 for N/a/b, at 2 for N/a/b…
  - Continue… Show on Board

Coded Formulation Example -- Factorials (Solution)

Case	Mav Ver	Target	T.O. Day
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	-1	1	1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1



- Note the geometric shape you built
- All variables are *orthogonal* (right angles)
- General algorithm to build -- vary A for half of runs, B for half of A and C for half of B …

Challenge 2: Systematically Search the Relevant Battlespace

4 reps 1 var

AGM-65D	AGM-65K
4	4

2 reps 2 vars

	AGM-65D	AGM-65K
Truck	2	2
Tank	2	2

1 reps 3 vars

		AGM-65D	AGM-65K
Eglin (Wet)	Truck	1	1
	Tank	1	1
Nellis (Dry)	Truck	1	1
	Tank	1	1

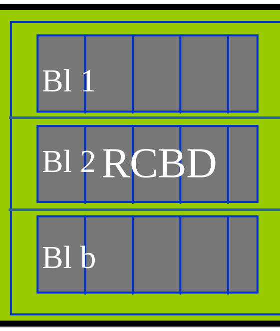
½ rep 4 vars

			AGM-65D	AGM-65K
Dawn (Transition)	Eglin (Wet)	Truck	1	
		Tank		1
	Nellis (Dry)	Truck		1
		Tank	1	
Midday (Stable)	Eglin (Wet)	Truck		1
		Tank	1	
	Nellis (Dry)	Truck	1	
		Tank		1

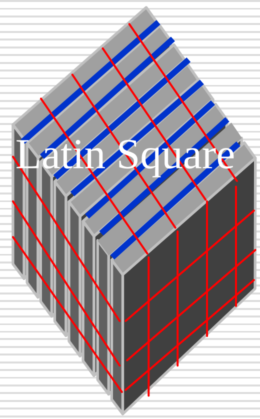
- Factorial (crossed) designs let us *learn more* from the same number of assets
- We can also use Factorials to *reduce assets* while maintaining confidence and power
- Or we can *combine* the two

All four Designs share the same ***power*** and ***confidence***

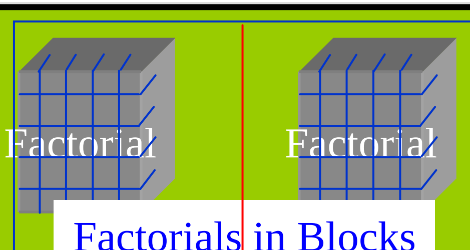
PLAN: With DOE we customize matrices from a collection of design templates



One Way With Blocks



M Way Designs



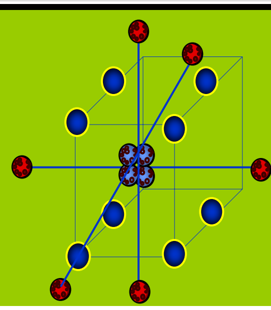
M Way Screening Designs

			High Altitude		Low Altitude	
ECM	Offset	Clutter	Auto Trk	Man Trk	Auto Trk	Man Trk
Wet	Short	High	12.1			5.6
		Low	14.7			4.6
	Long	High	12.2			7.8
		Low		13.6	3.4	

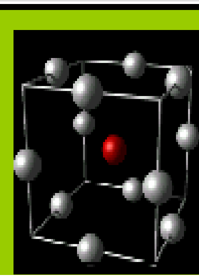
Fractional Factorials

(And Blocks)

M Way Response Surface Designs



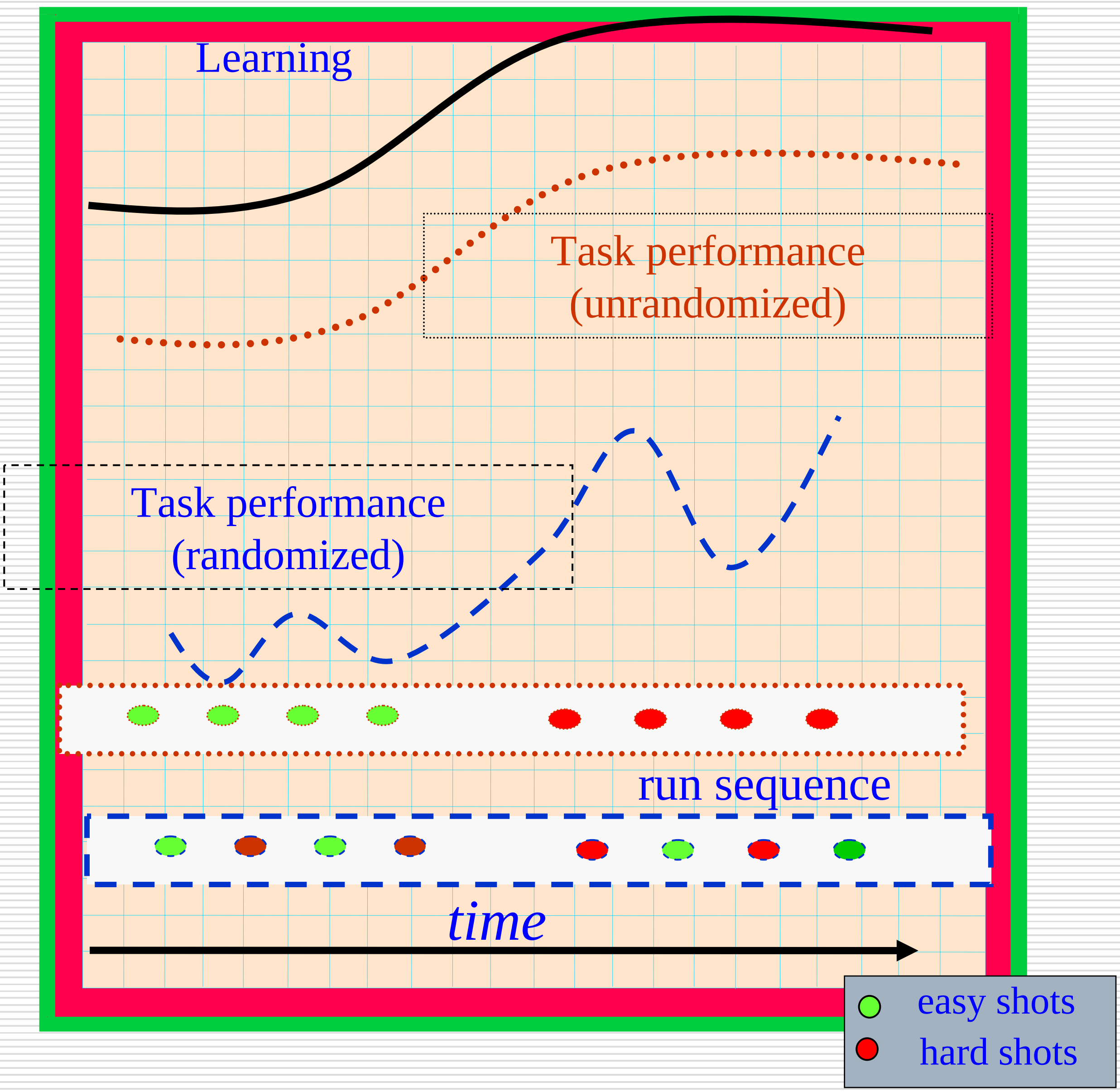
Central Composite



Box-Behnkin

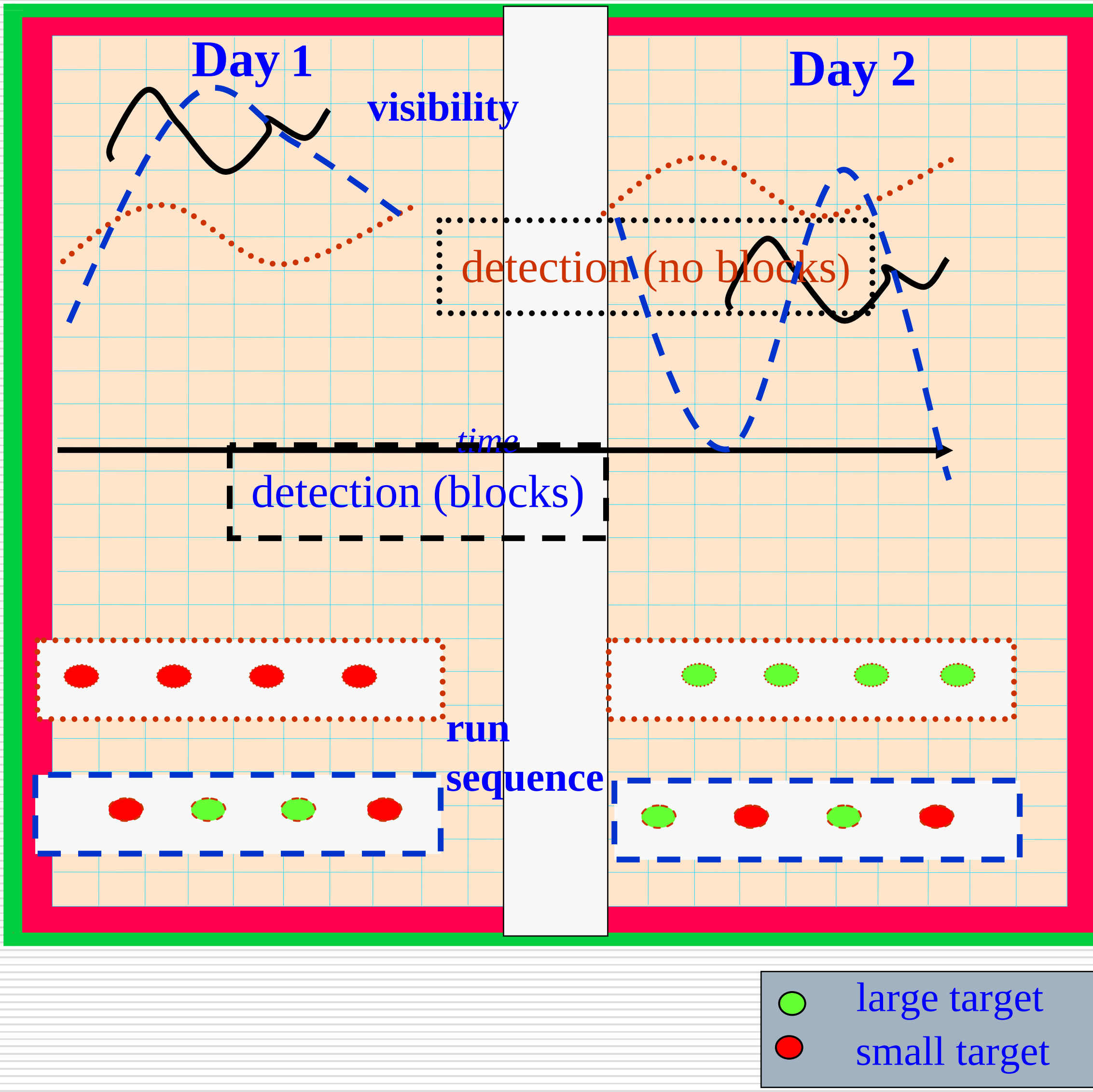
(And Blocks)

PRODUCE data with Random run order



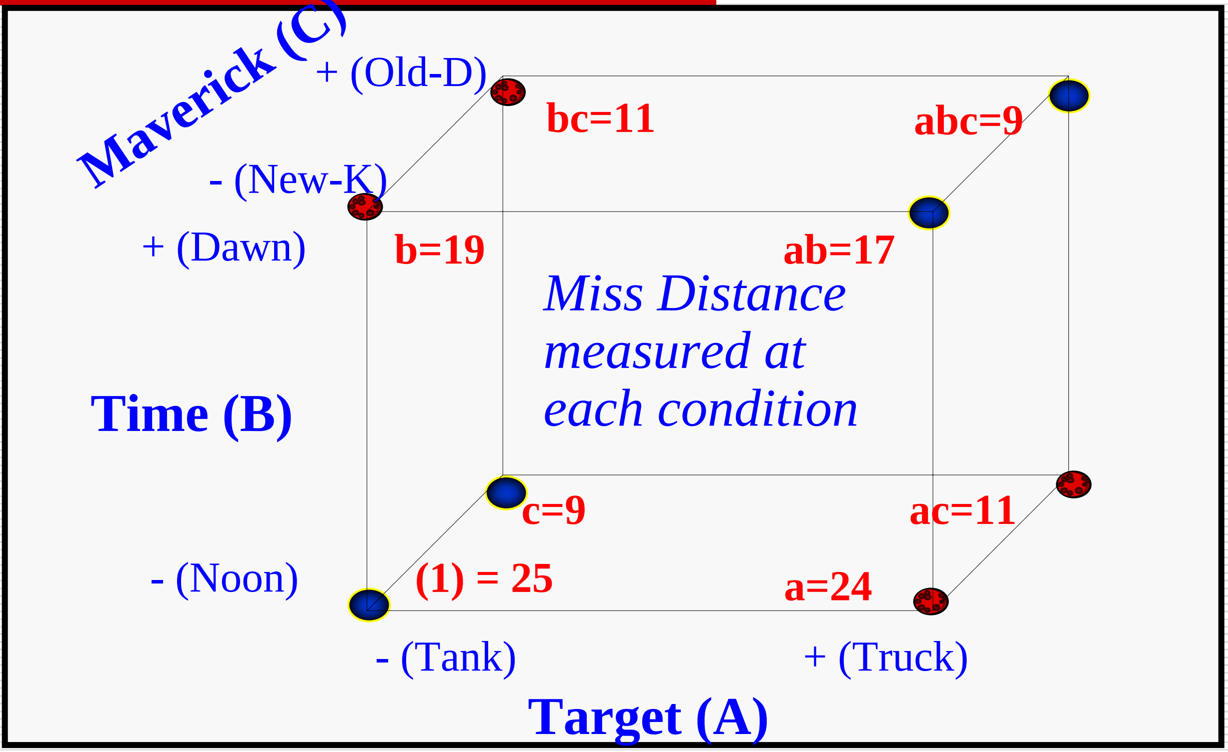
- **Randomizing** <u>runs</u> protects from unknown background changes **within** an experimental period (due to Fisher)

PRODUCE with Blocks



- **Blocking** <u>designs</u> protects from unknown background changes **between** experimental periods (also due to Fisher)

PONDER Simple Analysis -- How Conditions Affect MOP



- 8 Runs – Solve for 8 Unknowns:
- Overall Mean -- Effect (Target) -- Effect (Time) -- Effect (Maverick)
- + 3x2-ways, 3-way interactions

Designed Experiment *(n). Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).*

			Factorial Effect						
ase	Tr Combo	Mean	A	B	C	AB	AC	BC	ABC
1	(1)	1	-1	-1	-1	1	1	1	-1
2	c	1	-1	-1	1	1	-1	-1	1
3	b	1	-1	1	-1	-1	1	-1	1
4	bc	1	-1	1	1	-1	-1	1	-1
5	a	1	1	-1	-1	-1	-1	1	1
6	ac	1	1	-1	1	-1	1	-1	-1
7	ab	1	1	1	-1	1	-1	-1	-1
8	abc	1	1	1	1	1	1	1	1

Actual Maverick H/K Test – A great success

Design of Experiments DOE

Mission: MAV-1
Eglin Range   Range Time: 0700-0900 (Dawn)   Target: Point (Tank)   Weather: As Scheduled
Launch Airspeed: 400 KIAS

F-16 - #1   Left Wing: CATM-65K on LAU-88   Right Wing: TGM-65B on LAU-88

Run#	Target Type	Altitude	Sun Angle	Missile Type	Cueing	Comments*
1	Tank/Truck	18,000’	Sun at 6	H/K	Visual	
2	Tank/Truck	18,000’	Sun at 6	B	“	
3	Tank/Truck	1500’	Sun at 3/9	H/K	NAV/GPS	
4	Tank/Truck	1500’	Sun at 3/9	B	“	
5	Tank/Truck	1500’	Sun at 6	H/K	Radar	
6	Tank/Truck	1500’	Sun at 6	B	“	
7	Tank/Truck	18,000’	Sun at 3/9	H/K	LANTIRN	
8	Tank/Truck	18,000’	Sun at 3/9	B	“	

F-16 - #2   Left Wing: CATM-65K on LAU-88   Right Wing: TGM-65D on LAU-88

Run#	Target Type	Altitude	Sun Angle	Missile Type	Cueing	Comments*
1	Tank/Truck	18,000’	Sun at 6	H/K	Visual	
2	Tank/Truck	18,000’	Sun at 6	D	“	
3	Tank/Truck	1500’	Sun at 3/9	H/K	NAV/GPS	
4	Tank/Truck	1500’	Sun at 3/9	D	“	
5	Tank/Truck	1500’	Sun at 6	H/K	Radar	
6	Tank/Truck	1500’	Sun at 6	D	“	
7	Tank/Truck	18,000’	Sun at 3/9	H/K	LANTIRN	
8	Tank/Truck	18,000’	Sun at 3/9	D	“	

* After simulated pickle, simulate missile flyout

Typical Mav H/K F-16 Run Card

- Extensive captive carry
  - 22 sorties -Approx 100 sim shots
  - Old/new seekers on each wing to equalize Wx
  - 3 platforms: F-16, F-15E, A-10
  - Eglin & Nellis
- Results – approx 2x acq/trk range
- 9 shots comparable to current performance
- Type III Error – looking in the wrong place ...

A little bit of History …

- DOE has roots in Agriculture, Manufacturing
  - Royal Agricultural Experiment Station (R. Fisher -- 1920's)
  - Industrial Process Control (Deming, Shewart -- 1930's-40's)
  - Quality Revolution in Industry & Government (AFIT, NASA, Edwards, Ford, Motorola)
- Recent Developments include:
  - Fractional factorials (1960's)
  - Optimal Design of Experiments (1980's)
  - Measuring dispersion effects (1990's)
- People use DOE because it works …
  - Highly efficient
  - Unambiguous results
  - Long track record in science, engineering, medicine and industry
- DOE is:
  - A Test PROCESS Philosophy
  - A Collection of Test Matrix DESIGN Templates
  - A Set of ANALYSIS Methods

Historical DOE Timeline[1]

Hypothesis Tests (Gauss)

1800
1850
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000

Least Squares (Gauss, Legendre)

Regression Concepts (Person, Galton)

t-test (Gosset)

Factorial Experiments and ANOVA (Fisher)

Formalized Hypothesis Tests (Neyman, Pearson)

2^k Factorial Designs (Yates)

Fractional Factorial Designs (Finney, Rao)

2^{k-p} Fractional Factorial Resolution (Box, Hunter)

Taguchi develops his Methods

Central Composite Designs (Box, Wilson)

Optimal Designs (Kiefer, Wolfowitz)

Box-Behnken Designs (Box, Behnken)

Algorithm for D-optimal designs (Johnson, Nachtsheim)

Detecting dispersion effects ratio of variances (Montgomery)

Factorial Designs and ANOVA are DOE. DOE was first developed and used in crop trials by Sir R. A. Fisher, a mathematician and geneticist

1. Source: Appendix K -- *Understanding Industrial Designed Experiments,* Schmidt and Launsby,1998

Patron Saint of DOE – Sir R.A. Fisher

Sir Ronald Aylmer Fisher

http://www-history.mcs.st-andrews.ac.uk/history/Mathematicians/Fisher.html

Born: 17 Feb 1890 in London, England
Died: 29 July 1962 in Adelaide, Australia

Ronald Fisher received a B.A. in astronomy from Cambridge in 1912. There he studied the theory of errors under Stratton using Airy's manual on the Theory of Errors . It was Fisher's interest in the theory of errors in astronomical observations that eventually led him to investigate statistical problems.

Fisher gave up being a mathematics teacher in 1919 to work at the Rothamsted Agricultural Experiment Station where he worked as a biologist and made many contributions to both statistics and genetics. He had a long dispute with Pearson and he turned down a post under him, choosing to go to Rothamsted instead. There he studied the design of experiments by introducing the concept of randomisation and the analysis of variance, procedures now used throughout the world.

In 1921 he introduced the concept of likelihood. The likelihood of a parameter is proportional to the probability of the data and it gives a function which usually has a single maximum value, which he called the maximum likelihood.

In 1922 he gave a new definition of statistics. Its purpose was the reduction of data and he identified three fundamental problems. These are (i) specification of the kind of population that the data came from (ii) estimation and (iii) distribution.

The contributions Fisher made included the development of methods suitable for small samples, like those of Gosset, the discovery of the precise distributions of many sample statistics and the invention of analysis of variance. He introduced the term maximum likelihood and studied hypothesis testing.

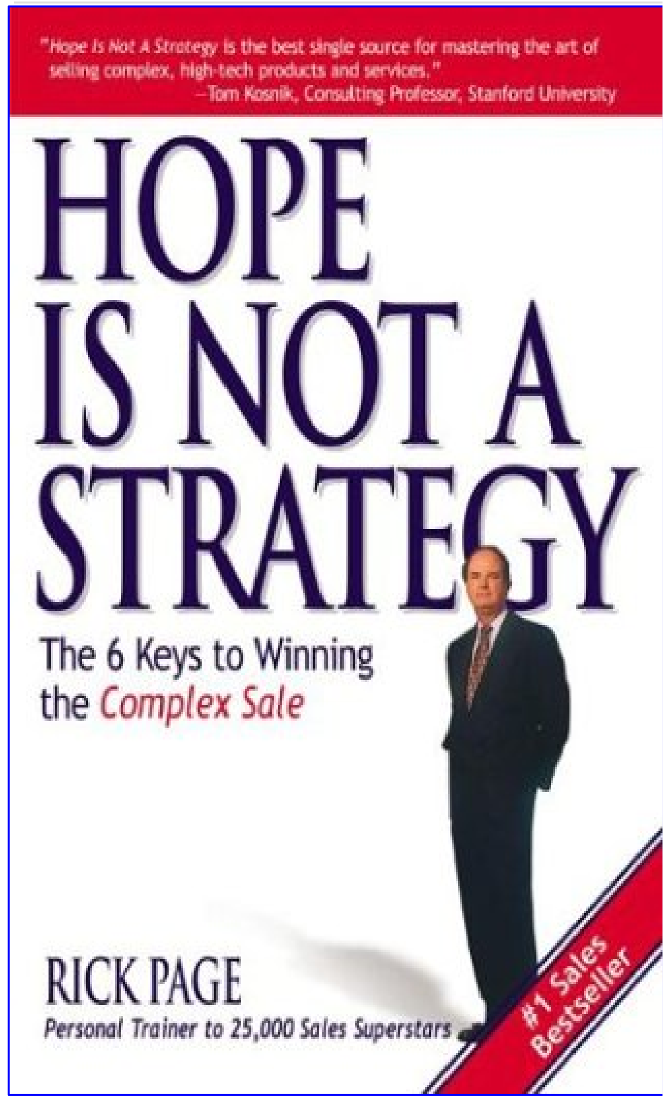
Fisher is considered one of the founders of modern statistics because of his many important contributions. He was elected a Fellow of the Royal Society in 1929, was awarded the Royal Medal of the Society in 1938 and he was awarded the Darwin Medal of the Society in 1948:- in recognition of his distinguished contributions to the theory of natural selection, the concept of its gene complex and the evolution of dominance. Then, in 1955, he was awarded the Copley Medal of the Royal Society:- in recognition of his numerous and distinguished contributions to developing the theory and application of statistics for making quantitative a vast field of biology.

He was awarded the Royal Medal of the Royal Society in 1938 and the Copley Medal in 1955. You can see a history of the Royal Medal and a list of the winners in our archive and a history of the Copley Medal and a list of the winners.

A Strategy to be the *Best* … In Test Design

- Train 53d Wing Leadership in statistical thinking for Test
- Adopt the most powerful test strategy (DOE)
- Train total team from PM to OA
- Develop mentors and formally mentor test teams
- Revise Wing procedures
- Share these test improvements

Adopters	**Targets**
■ AFOTEC	■ HQ AFMC
■ 18 FTS (AFSOC)	■ AAC
■ F-15 OFP CTF	■ 46 TW
■ USAF TPS	
■ AFFTC	



The *Least* You Should Recall…

- The Central Problem of Test is inferring what the real world is like based on our sample
  - We can make one of two errors – false positive and false negative
  - Want to design our tests to minimize these risks
- A Designed Experiment is about the Design – the pattern of test conditions we run
  - Two variables are Confounded when we cannot separate their effects on the MOP (response)
  - DOE avoids **confounding** two or more variables
- Basic process to Construct is Process-Plan-Produce-Ponder
- DOE has 80-year track record in all areas of science, engineering, and test
- DOE is a test strategy in the test big picture
- DOE has made 53rd Wing tests better, faster and cheaper
- Resampling (Monte Carlo) statistics give us a powerful tool to directly attack hard problems

Resources

- Links: https://wwwmil.53wg.eglin.af.mil/milweb/index.htm -- *Training* menu (53d Wing Courses)
- www.statsoft.com (Statistica)
- www.minitab.com (MiniTab)
- www.statease.com (Design Ease)
- www.resample.com (Resampling Stats)
- Books:
  - *Design and Analysis of Experiments*, 5th Ed Douglas Montgomery, 2001
  - *Statistics for Experimenters*, Box, Hunter & Hunter, 1978
  - *Design of Experiments: Statistical Principles of Research Design and Analysis,* Robert Kuehl

Matrix of Correlation Coefficients among X Variables					
	A	B	C	A	B
A	1.00				
B	-0.15	1.00			
C	-0.26	-0.27	1.00		
D	-0.39	-0.07	0.07	1.00	
E	-0.24	-0.04	-0.43	0.62	1.00

Session 2 – Comparing Test Strategies

- Is DOE the Best Test Strategy?
- Other Strategies
  - One factor at a time
  - Previous experience
  - Scenario testing
- Interactions
- Advantages of DOE

What's *Best*? Many methods of test have been tried

- Our next example will *characterize* the *B-1B Radar Target Location Error* as a function of several factors

Altitude
Target RCS
Angle off nose
Aircraft tail number
Target Range
Radar calibration date
Last doppler update
Target Elevation
Operator Skill Level

Inputs (test conditions)

Process: Radar Ground Mapping

Target location accuracy

Outputs (MOPs)

- **Previous experience**
- **Contractor generated**
- **Intuition**
- **One factor at a time (OFAT)**
- **Scenario (best guess or case)**

B-1 Radar Target Location Error Test Space Magnitude

- How big is the test space?
- Consider 9 factors minimum - two levels each
  - For example, set angle off nose to 15 or 30 degrees
- If we tested each possible combination, how many would there be?

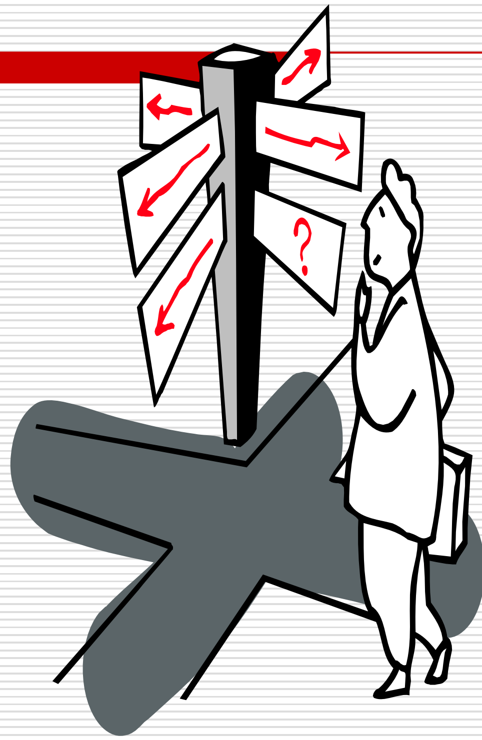
$$2^9 = 512$$

How should we examine this test space?

BTW … the "Plus 2 Ways"

- Previous experience
  - Picture a young 1Lt assigned as Lead Analyst, F-16C Block 0
  - Q: What will he do?
  - A: The best he can …
  - Graybeard advice – get the F-16A MOT&E Test Plan!
  - Obviously – not a general solution … how good is F-16A Test? What about Global Hawk and JSF?
- Run contractor-selected test points
  - Current 53d Wing CTF Test – run these 227 cases 3X each
  - Q: Why these points? Why 3 times?
  - A: Because that's what we did last time. We do 3 to get 1.
  - Q: How do you compare new OFP to old OFP performance?
  - A: Mostly we just look at them.
- When we look under covers … mostly one of three strategies

Intuition



- One way to examine the test space is to rely on intuition
- Intuition can greatly benefit a test; however…
  - Requires deep subject matter knowledge that is not always available
  - Typically only "discover" what you already believe to be true
  - Lacks proof
- Intuition can assist with test design but it should never be the sole strategy

OFAT Design looks like

The Variables (Factors)

A. Angle off nose
B. Range to target
C. Clutter background
D. Target RCS
E. Aircraft velocity
F. Aircraft altitude
G. Years OSO experience
H. Time since radar calibration
J. Time since doppler update

Total number of combinations = ?

At two levels each variable we have 2x2x2… = 2^9 = 512 combinations

110 total resources used

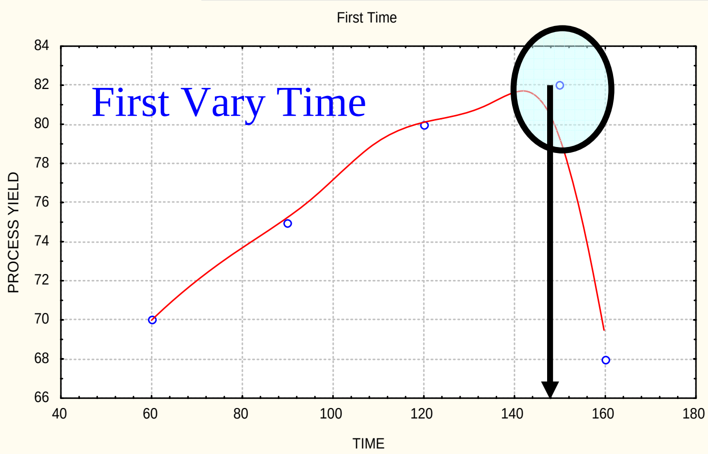
Case	A	B	C	D	E	F	G	H	J		Y1	…	Y10	Avg
1	low	low	low	low	low	low	low	low	low					
2	hi	low	low	low	low	low	low	low	low					
3	best	hi	low	low	low	low	low	low	low					
4		best	hi	low	low	low	low	low	low					
5			best	hi	low	low	low	low	low					
6				best	hi	low	low	low	low					
7					best	hi	low	low	low					
8						best	hi	low	low					
9							best	hi	low					
10								best	hi					
11									best					

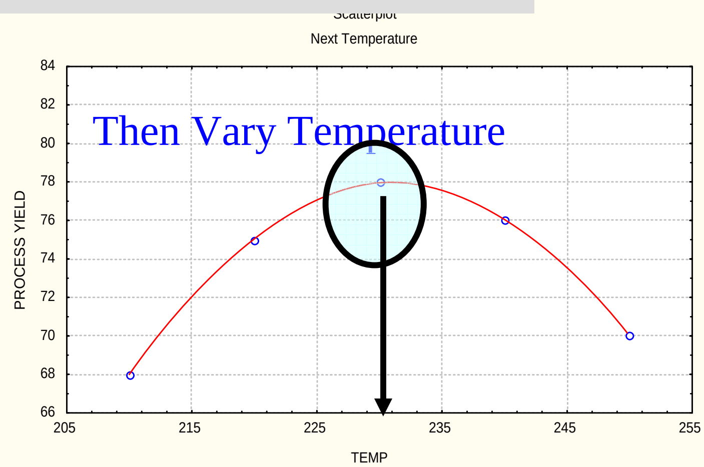
We're going to run 10 because we're 10-guys and that's what we do.

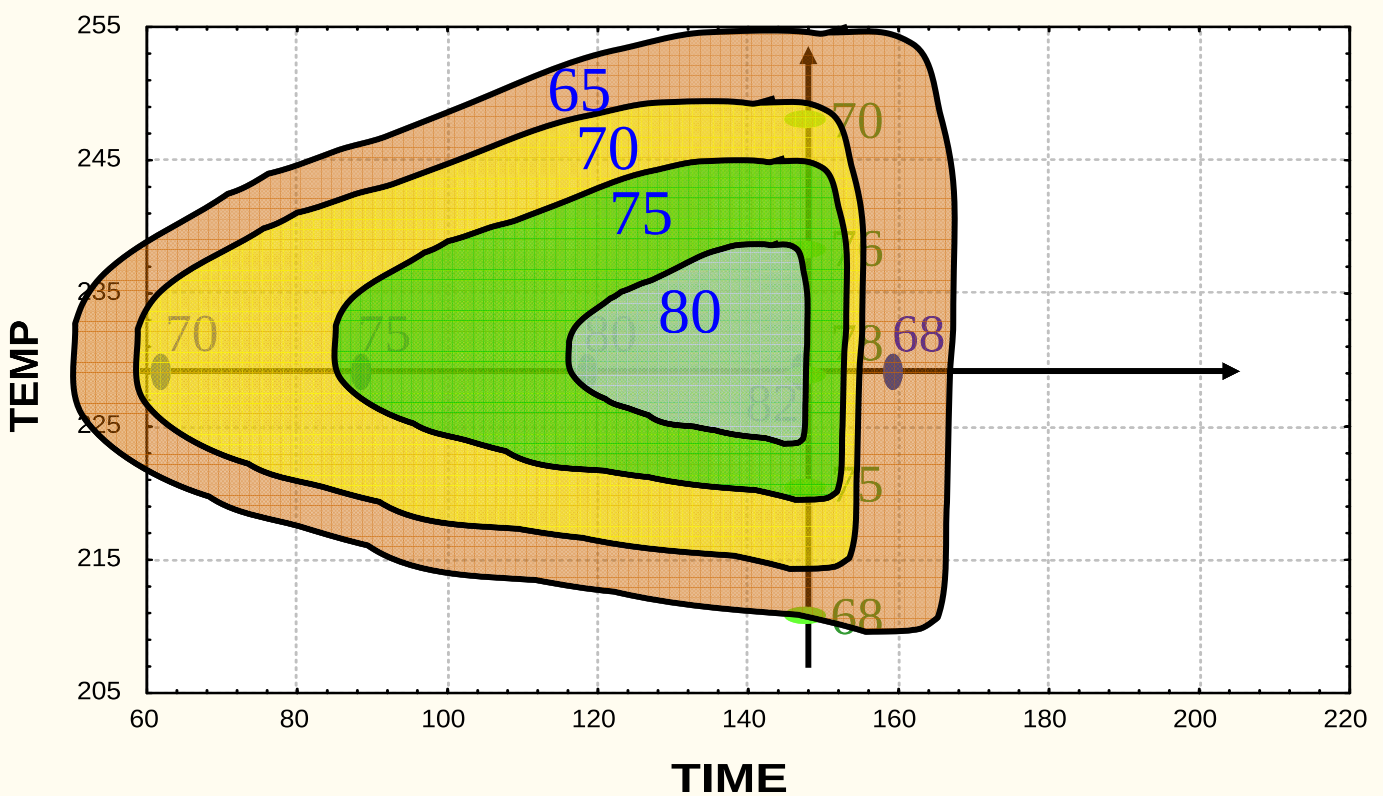
- 11 combinations tested
- No math model to predict other 501 combos
- No ability to estimate if variables ***interact***

Yesterday’s typical OFAT Test – Brewing Beer …

Objective is to maximize “beer goodness”







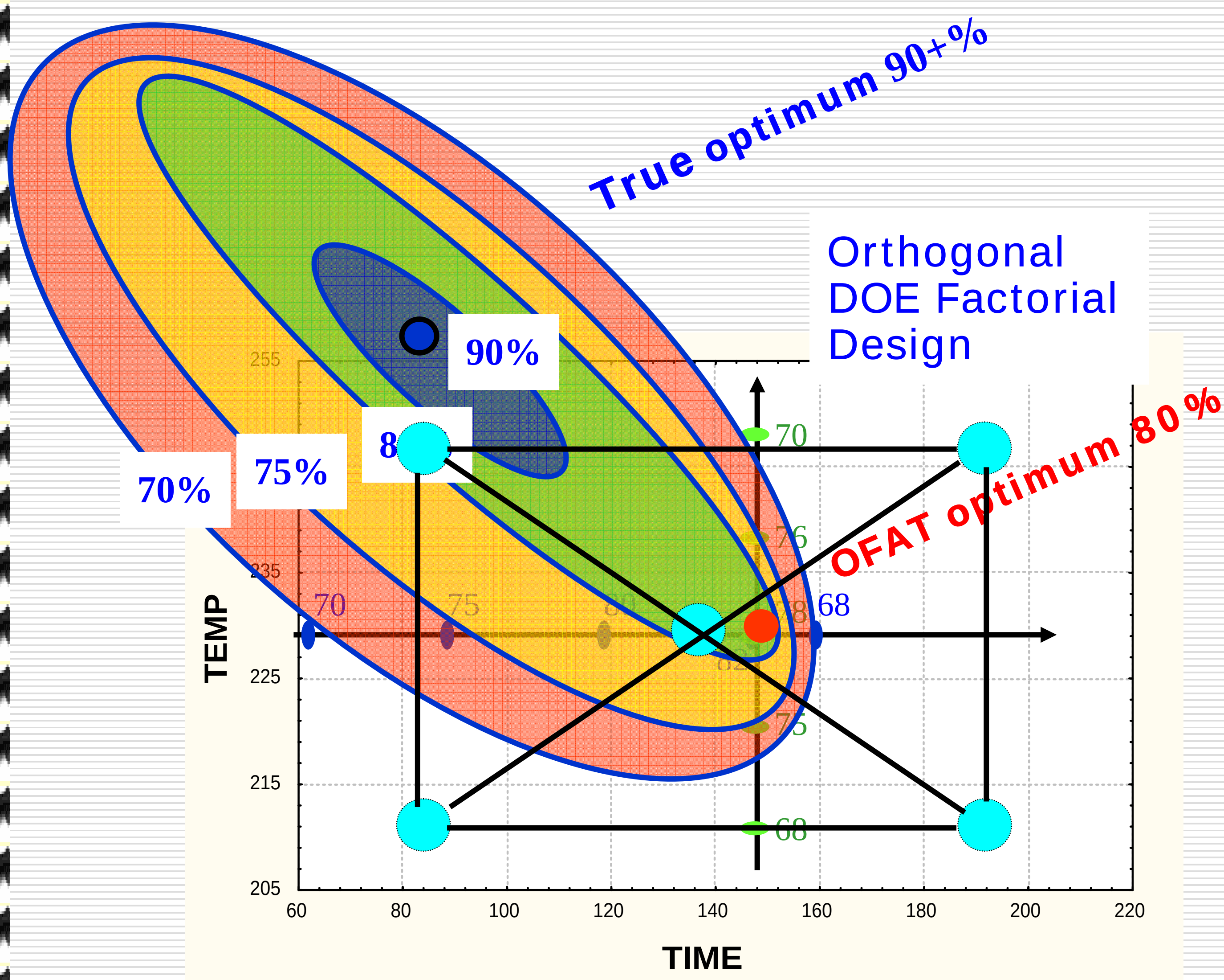
- OFAT works if response contours align with axes… but what if the contours are **not** aligned?

OFAT Assumption ...

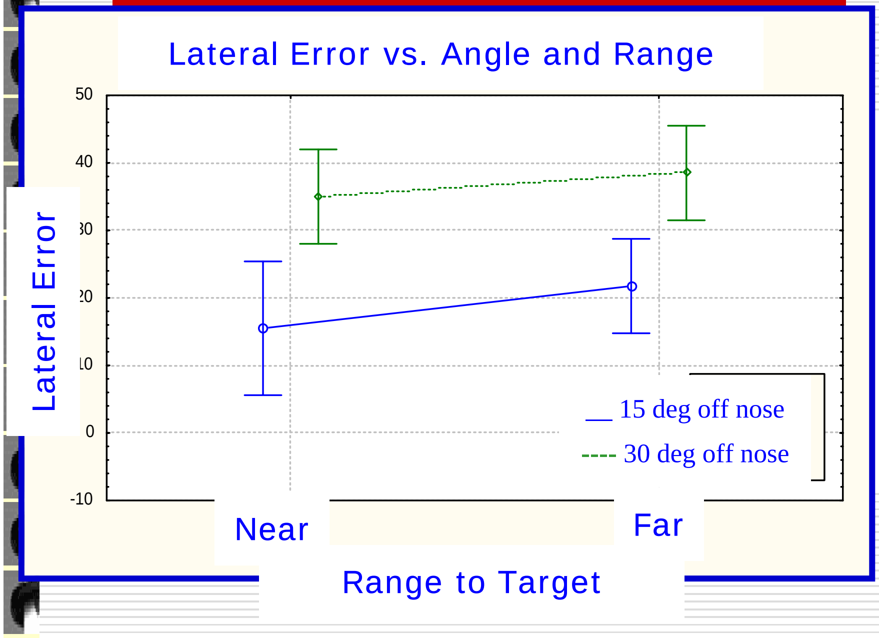
If contours are not aligned with axes, we miss the optimum.

Problem is not simply to find best time and then best temperature -- we must find if the variables **interact**

(Response contours are often ridges, saddles and other shapes at an angle to our control axes)



Back to B-1 Radar: the meaning of interaction



Parallel slopes of A with B -- response at level of B does not depend on level of A

Also -- can talk about variable settings independently -- “15 degrees off nose always best”

Simple model:

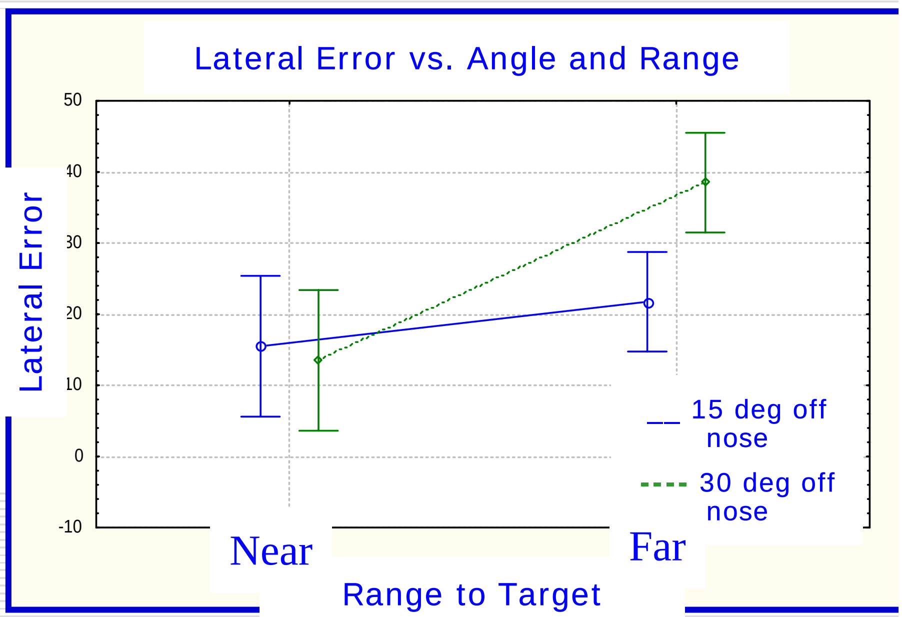
$Y=b_0+b_1*Range+b_2*Angle$

Nonparallel slopes of A with B -- response at level of B <u>does</u> depend on level of A

Also -- must talk about variable settings together -- “Long range targets detected best at 15 degrees off nose”

More complex model with interaction term:

$Y=b_0+b_1*$ Range $+b_2*$ Angle $+$ $\boldsymbol{b_3 * Range * Angle}$



OFAT cannot detect interaction

Best Guess or “Cases” Testing relies on intuition (not fact)

Case	A	B	C	D	E	F	G	H	J		Y1	…	Y17	Avg
1	low	hi	low	hi	low	hi	low	low	low					
2	hi	hi	low	low	low	low	low	low	low					
3	low	low	hi	low	hi	low	low	low	low					
4	low	low	low	low	low	hi	low	low	low					
5	low	low	low	low	low	low	low	hi	hi					
6	low	low	low	low	low	hi	hi	low	low					
7	low	low	low	low	low	low	low	low	low					
8	low	low	hi	hi	hi	low	low	low	low					
9	low	hi	low	low	hi	low	low	low	low					
10	low	low												
11	hi	hi												

Best guess means choosing those combinations the subject expert feels most likely contain the answer. Usually, the organization has a magic number of replications (3,8,30…) they believe will be “good” in some unspecified sense. Often they say “significant” or “data is normal.”

Case	A	B	C	D	E	Ybar
1	low	hi	hi	low	low	20
2	hi	low	low	hi	low	9
3	low	hi	hi	low	low	3
4	low	low	low	low	low	8
5	hi	hi	low	low	low	7
6	low	hi	hi	low	low	8
7	low	low	hi	hi	hi	24
8	low	low	hi	low	low	16
9	low	hi	low	hi	low	2
10	hi	hi	hi	hi	hi	30
11	low	low	low	low	low	6

If runs 1-11 were best guess, what can we conclude?

1. Does factor E affect the average?
2. Can we separate effects of B and C?
3. What D, E combo is missing?

Scenario Designs - Confounding among Predictors

Matrix of Correlation Coefficients among X Variables

	A	B	C	A	B
A	1.00				
B	-0.15	1.00			
C	-0.26	-0.27	1.00		
D	-0.39	-0.07	0.07	1.00	
E	-0.24	-0.04	-0.43	0.62	1.00

- Such correlations make *independent* estimates of effects impossible
- In English – the effects of the test conditions are muddled: low effects are magnified and high effects are reduced
- And, we cannot detect when two conditions work together to change the response (*interactions*)

Maverick example-DOE *Crosses* factors to avoid confounding and estimates effects and interactions

4 reps 1 var

AGM-65D	AGM-65K
4	4

2 reps 2 vars

	AGM-65D	AGM-65K
Truck	2	2
Tank	2	2

1 reps 3 vars

		AGM-65D	AGM-65K
Eglin (Wet)	Truck	1	1
	Tank	1	1
Nellis (Dry)	Truck	1	1
	Tank	1	1

½ rep 4 vars

			AGM-65D	AGM-65K
Dawn (Transition)	Eglin (Wet)	Truck	1	
		Tank		1
	Nellis (Dry)	Truck		1
		Tank	1	
Midday (Stable)	Eglin (Wet)	Truck		1
		Tank	1	
	Nellis (Dry)	Truck	1	
		Tank		1

- Factorial (crossed) designs let us *learn more* from the same number of assets
- We can also use Factorials to *reduce assets* while maintaining confidence and power
- Or we can *combine* the two

All four designs share the same ***power*** and ***confidence***

DOE Screening Design vs. OFAT

Case	A	B	C	D	E	F	G	H	J		Ybar
1	low	low	low	low	low	hi	hi	hi	hi		
2	hi	low	low	low	low	hi	low	low	low		
3	low	hi	low	low	low	low	hi	low	low		
4	hi	hi	low	low	low	low	low	hi	hi		
5	low	low	hi	low	low	low	low	hi	low		
6	hi	low	hi	low	low	low	hi	low	hi		
7	low	hi	hi	low	low	hi	low	low	hi		
8	hi	hi	hi	low	low	hi	hi	hi	low		
9	low	low	low	hi	low	low	low	low	hi		
10	hi	low	low	hi	low	low	hi	hi	low		
11	low	hi	low	hi	low	hi	low	hi	low		
12	hi	hi	low	hi	low	hi	[illegible]	[illegible]	[illegible]		
13	low	low	[illegible]	[illegible]	[illegible]	hi	hi	low	low		
14	hi	low	hi	hi	low	hi	low	hi	hi		
15	low	hi	hi	hi	low	low	hi	hi	hi		
16	hi	hi	hi	hi	low	low	low	low	low		
17	low	low	low	low	hi	low	low	low	low		
18	hi	low	low	low	hi	low	hi	hi	hi		
19	low	hi	low	low	hi	hi	low	hi	hi		
20	hi	hi	low	low	hi	hi	hi	low	low		
21	low	low	hi	low	hi	hi	hi	low	hi		

DOE Fractional Factorial

In 32 samples, with this fractional factorial ((1/16th fraction of resolution IV), we can estimate the main effects of 9 variables. Can collapse across one or more variables to estimate spread. If 4 of 9 factors are not active, we have all combinations in any remaining 5 variables.

DOE uses 29% of OFAT samples

Case	A	B	C	D	E	F	G	H	J		Y1	…	Y10	Avg
1	low	low	low	low	low	low	low	low	low					
2	hi	low	low	low	low	low	low	low	low					
3	best	hi	low	low	low	low	low	low	low					
4		best	hi	low	low	low	low	low	low					
5			best	hi	low	low	low	low	low					
6				best	hi	low	low	low	low					
7					best	hi	low	low	low					
8						best	hi	low	low					
9							best	hi	low					
10								best	hi					
11									best					

OFAT Design

We're going to run 10 because we're 10-guys and that's what we do.

DOE Screening Design vs. Scenario Testing

Design of Experiments

DOE

Case	O	T	N	R	A	T1
1	low	low	low		low	
2	hi	low			low	
3	low	hi	low	low	low	
4	hi	hi	low	low	low	
5	low	low	hi	low	low	
6					low	
7	low	hi	hi	low	low	
8	hi	hi	hi	low	low	
9						
10						
11	low	hi	low	hi		
12	hi	hi	low	hi		
13	low	low	hi	hi		
14	hi	low	hi	hi		
15	low	hi	hi	hi		
16	hi	hi	hi	hi		

DOE Fractional Factorial

16 total events

In 16 events, with this *fractional factorial matrix* ((1/2 fraction of resolution V), testers can estimate the effects of each test condition. Can collapse across one or more variables to estimate variance. If 1 of 5 conditions don't affect the MOP, matrix has all combinations of the remaining 4 variables.

DOE uses 20% of "Best Guess" samples

Event	OSO					
1	Jones					
2	Smith					
3	Jones	2	2145	12	10k	3
4	Jones	1	168	12	10k	8
5	Smith	2	168	12	10k	7
6	Jones	2	2145	12	10k	8
7	Jones	1	2145	34	20k	24
8	Jones	1	2145	12	10k	16
9	Jones	2	168	34	10k	2
10	Smith	2	2145	34	20k	30
11	Jones	1	168	12	10k	6

B-1 Radar Mapping in One Mission (28th TS)

Problem: Characterize B-1 radar coordinate accuracy for variety of operational setups.

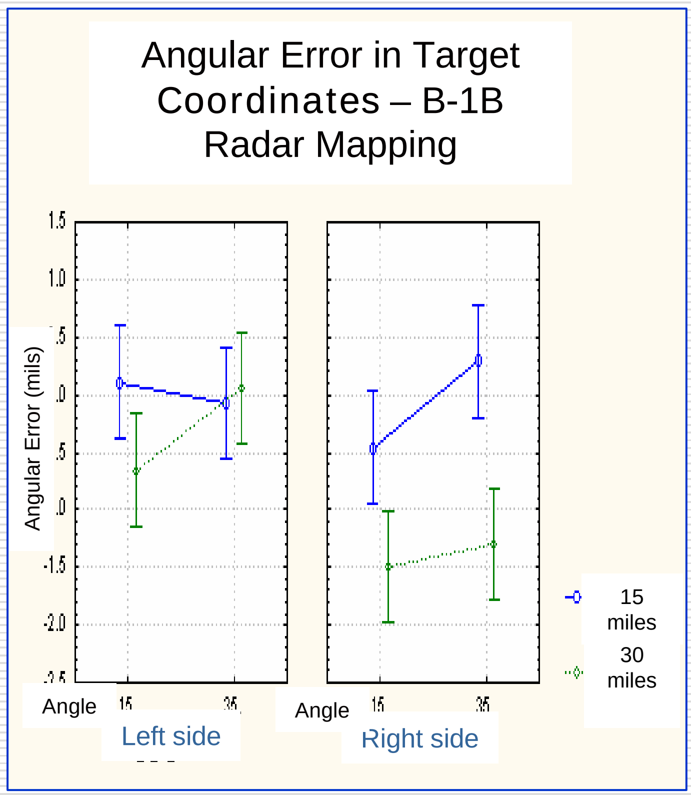
Design:

Response: absolute error

Conditions: angle, side of nose, tail number, target, and range to target.

4 replicates.

Result: Similar accuracy across scan volume, target type, tail number.



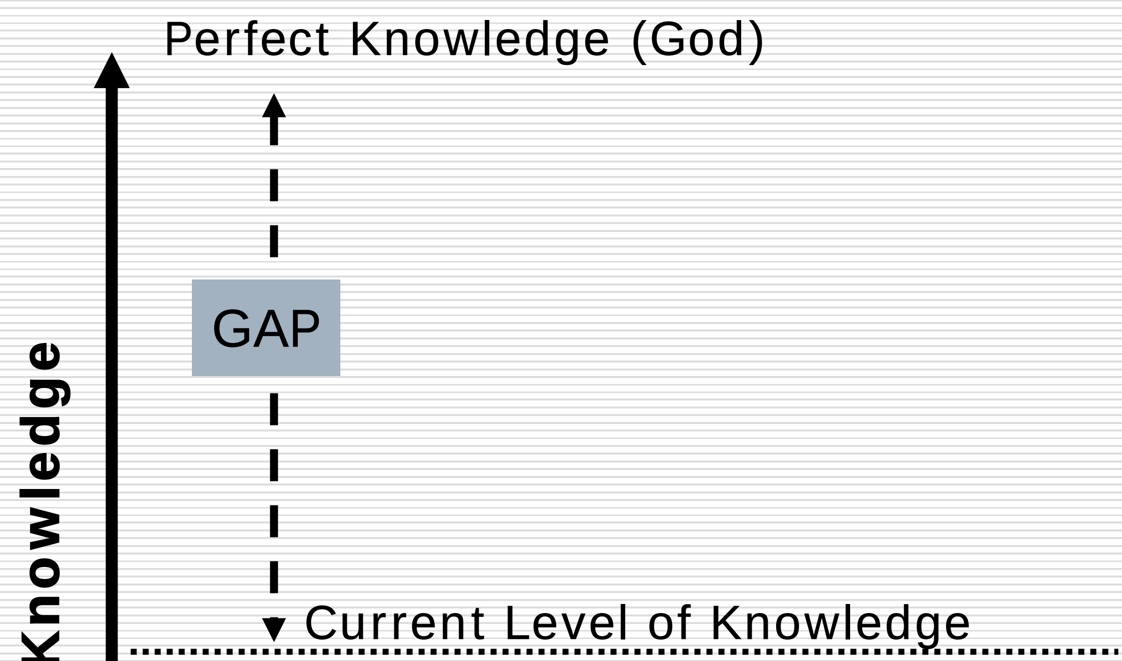
Result: Single two-aircraft mission answered accuracy questions raised by 7 previous missions using conventional test methods.

Other Recent Project highlights …

- 2X increase in Predator Hellfire firing envelope in a week's testing
- 50% reduction in expenditures and runs required for B-1 Block E upgrade test with increased data fidelity
- 45% reduction in sorties and helped define test objectives and procedures for $325K HTS verification
- 33% reduction in bombs required during evaluation of 4 weapon types off 10 B-1 sorties
- Uncovered fundamental TTPs in multi-platform E-CAS TD&E in 40 sorties
- 25% reduction in resources using screening method for Chemical Detector test
- $36K in lab costs saved for $144K Expeditionary-Deployable Oxygen Concentrator System (EDOCS) test
- 133 passes to 62 in B-1 PFS 4.2 ECM systems upgrade

We have more than 150 DOE test projects across full range of ACC systems

Why DOE -- Postscript



We use DOE to interrogate the process and improve our knowledge of how our process works. The goal is a systematic method to efficiently and unambiguously improve our outcomes.

Compared to any other systematic method, DOE designs:

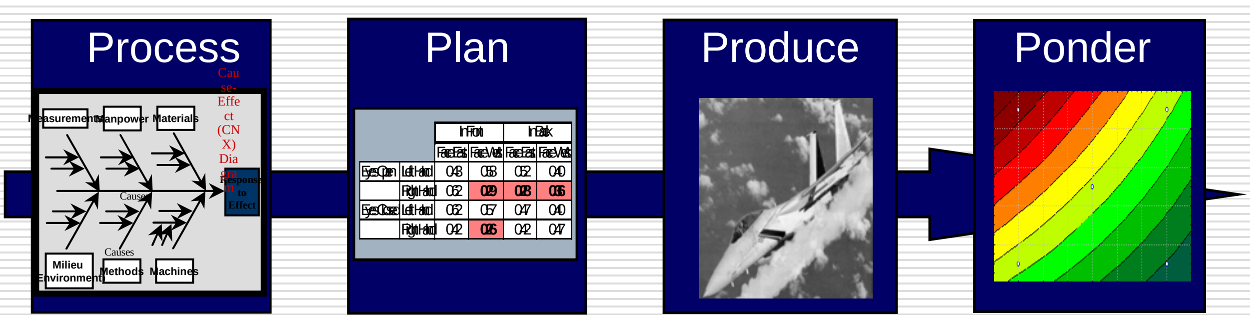
- –Yield **Better** process understanding
- –Can be planned and analyzed **Faster**
- –Are **Cheaper,** using only 20-80% of usual resources

The *Least* You Should Recall…

- Three methods (or strategies) of test are commonly used – OFAT, scenario, and intuition
- Intuition is important, but not sufficient as our sole test strategy
- OFAT, by its nature, cannot detect interactions – and interactions are common
- The OFAT search vector leaves large areas unexplored
- Two variables interact when the setting of one variable changes the effect on the response of the second variable
- Scenario designs contain complex confounding patterns making cause-and-effect reasoning difficult
- Scenario designs afford spotty coverage of the total factor space
- DOE designs efficiently and effectively tell us how our operations work

Session 3 – Four Steps to Design

- DOE Spiral 2 – SAM-in-a-Box Part Deux
  - Decomposing the PROCESS
  - Developing the test PLAN
  - PRODUCE the observations
  - PONDER what we have learned
  - Loop … and iterate

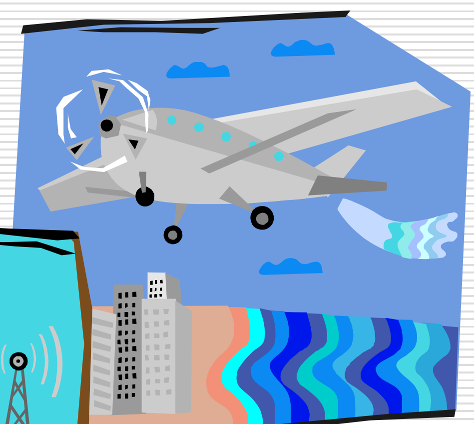


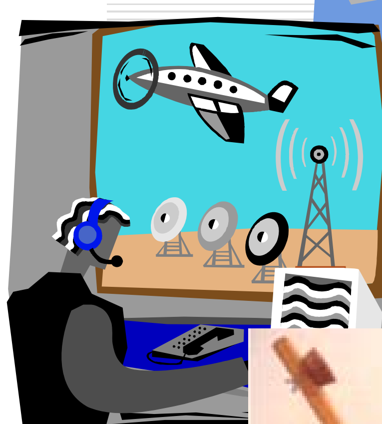
Break – Spiral 2 Back to the SAM-in-a-Box

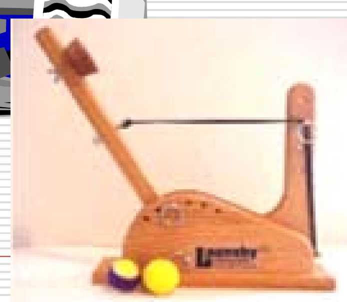
- We'll do a simplified multi-factor design and simplified ANOVA

News Flash! US Company Indicted for Iraq Arms Sales!

- The SAM-in-a-Box is a potent threat!
- Sold to Saddam by Stat-a-Pult Inc.
- We must conduct an exploitation HWIL simulation today
- MOP – Launch Range
- Conditions – projectile, propellant, Tgt RCS & velocity, ECM ... etc

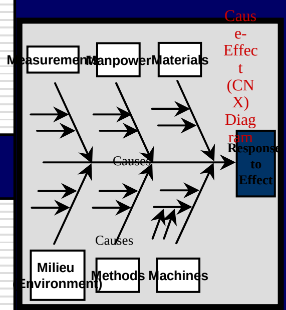




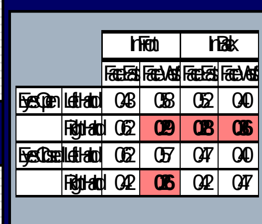


Spiral 1 --How We Design Tests in 4 Stages

Process

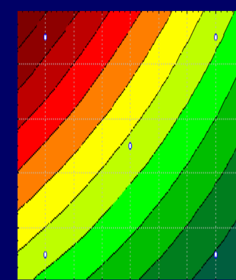


Plan



Produce

Ponder



I Project Description and Decomposition

II Plan the Test Matrix

III Produce the Observations

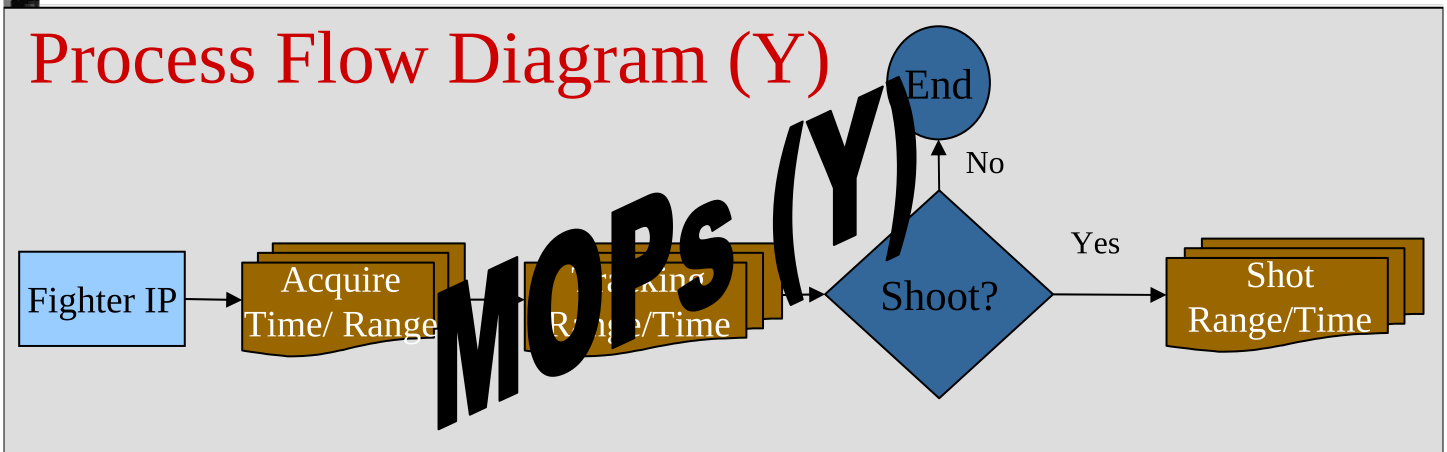
IV Ponder the Results

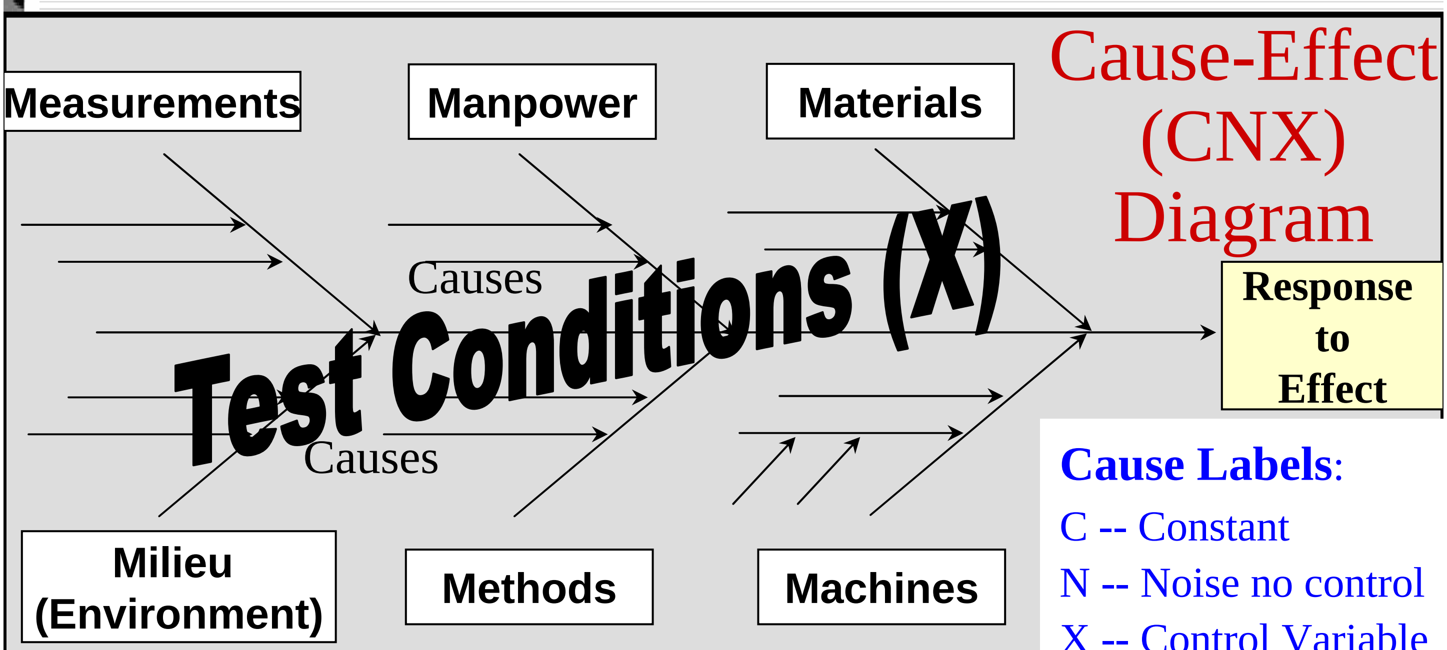
We'll stroll quickly through examples of each step for Maverick Problem

Stage I Samples of Various Test Objectives

- Compare system to a fixed standard (ORD, specification, demo goal)
- Compare system to older version (OFP or software/hardware upgrade, new mission data, etc.)
- Characterize system performance (test first, then evaluate later)
- Optimize the performance of the system under test (maximize or minimize)
- Make system most robust to environmental conditions (weather, threat actions, countermeasures, operator experience)
- Minimize the variability of system performance
- Troubleshoot faulty system performance (false alarms)

Two Tools to Turn Eng / Ops Art into Science





Cause Labels:

C -- Constant

N -- Noise no control

X -- Control Variable

Tool 1 –Process Flow for Table of RESPONSES (MOPs)

Process Flow Diagram

Table of MOPs

N	Variable	Units	Range	Priority	Data Elem	Source

Tool 2 – Fishbone Table of TEST CONDITIONS

Fishbone or CNX Diagram

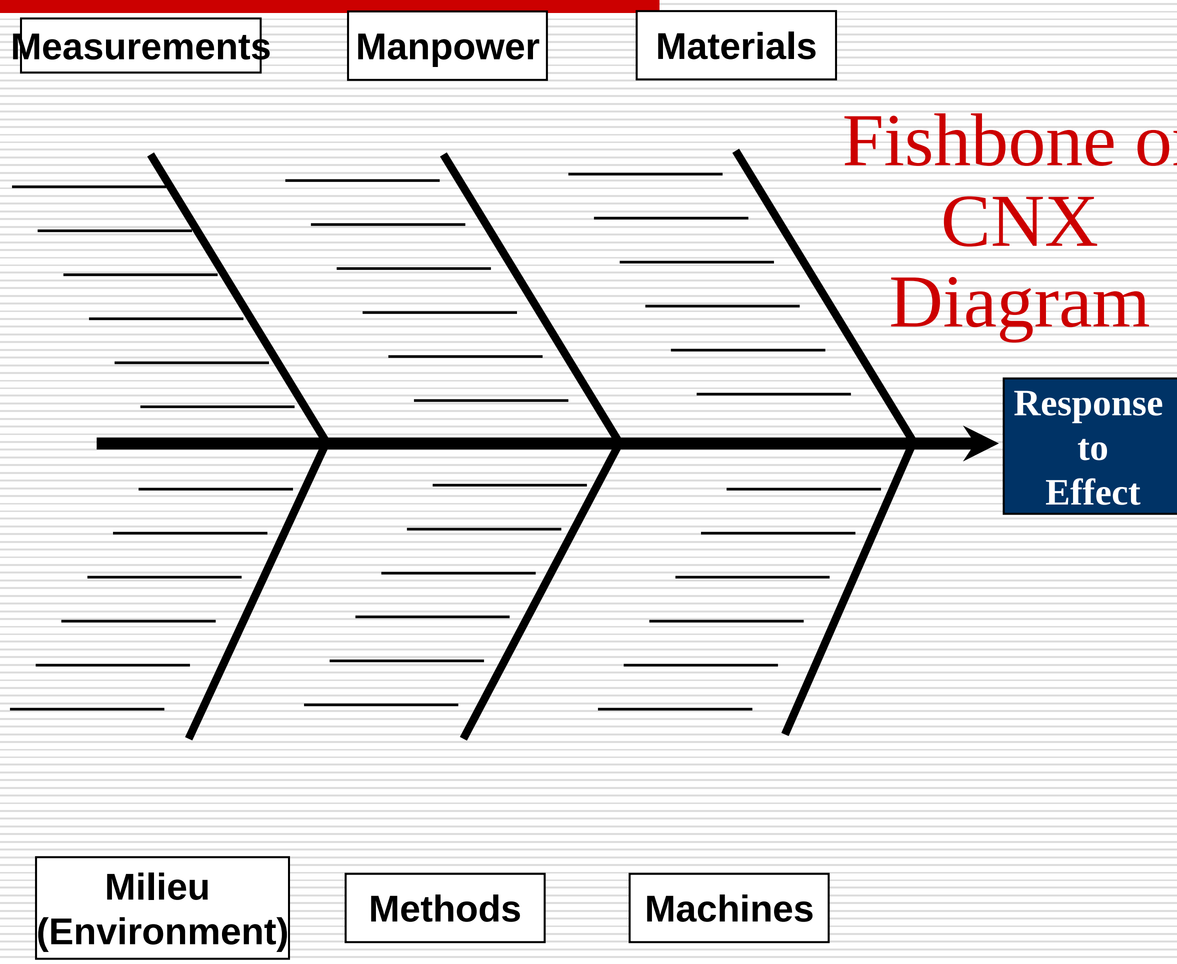
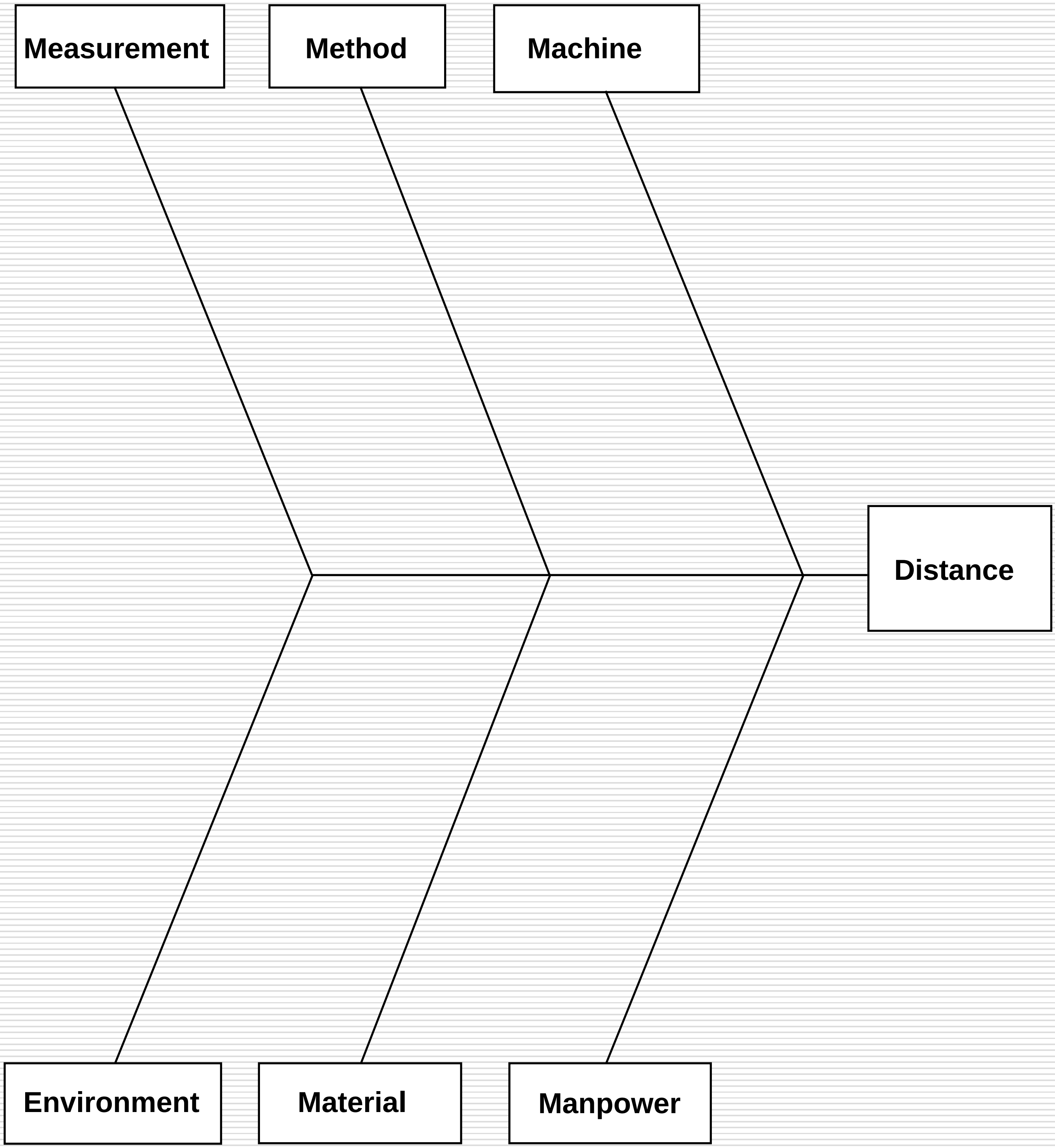


Table of Test Conditions

N	Variable	Units	Range	Priority	Exp. Control	Design Range

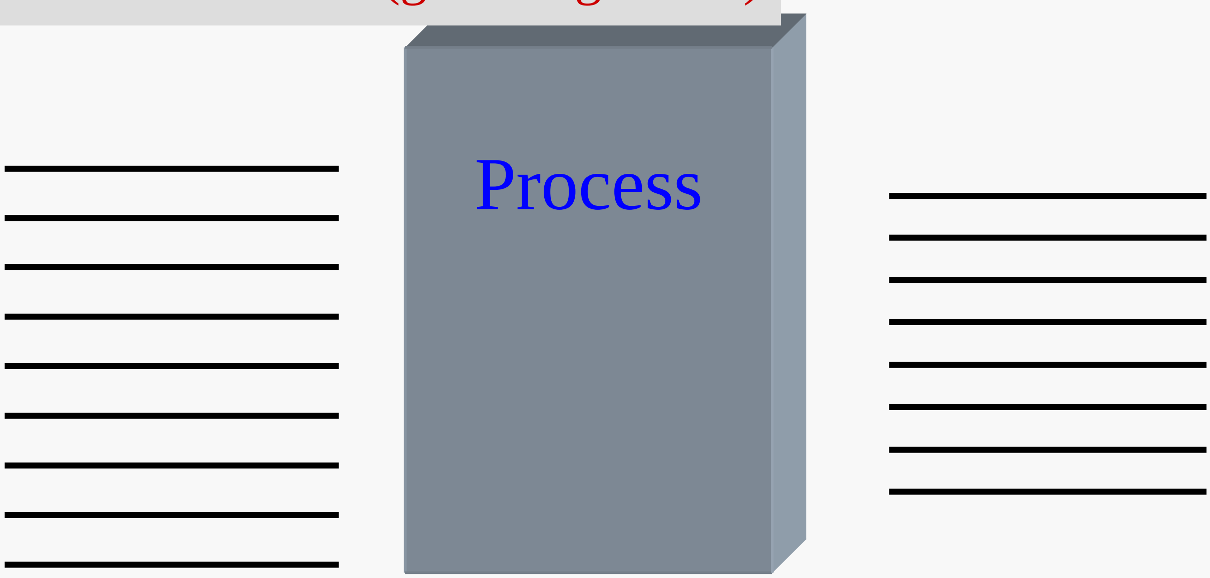
Cause and Effect Diagram



BTW When Time is Short -- Tool 3 Black Box Diagram

Process Flow Chart

Process Black Box (gozinta-gozouta)



Inputs (Test Conditions)

Outputs (MOPs)

Design Algorithm for Full Factorial

Case	Stop	Tgt	Ball	PullBack
1	1	F-16	Red	135
2	1	F-16	Red	180
3	1	F-16	Yellow	
4	1	F-16	Yellow	
5	1	F-117		
6	1	F-117		
7	1	F-117		
8	1	F-117		
9	3			
10	3			
000	000			
23	5			
24	5			

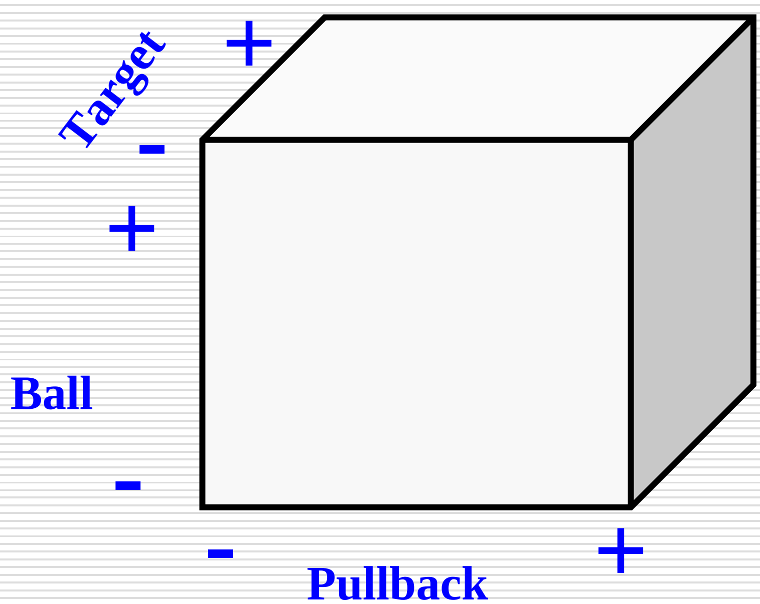
- How many Variables (A,B,C … K)?
- How many settings (levels) for each Variable (a, b, c, … k)?
- How many runs total?
  - a *b* … *k
  - Example: 4 vars: 1 at 3, 3 at 2 levels
  - Runs (N) = $3*2*2*2 = 3 * 2^3$ = 24 runs (test cells)
- Construct Design --
  - Var A set at 1 for N/a, level 2, etc.
  - With A at 1, set B at 1 for N/a/b, at 2 for N/a/b…
  - Continue… Show on Board

PLAN: Create Design

- Let each table take 20 minutes to formulate a design. We will discuss your results as a group.
- Consider the test event under consideration
- Use only **three** variables at any number of levels each -- instructor will supply them.
- Use instructor-supplied measurement as your MOE
- How many total runs for one replicate?
- Justify your design

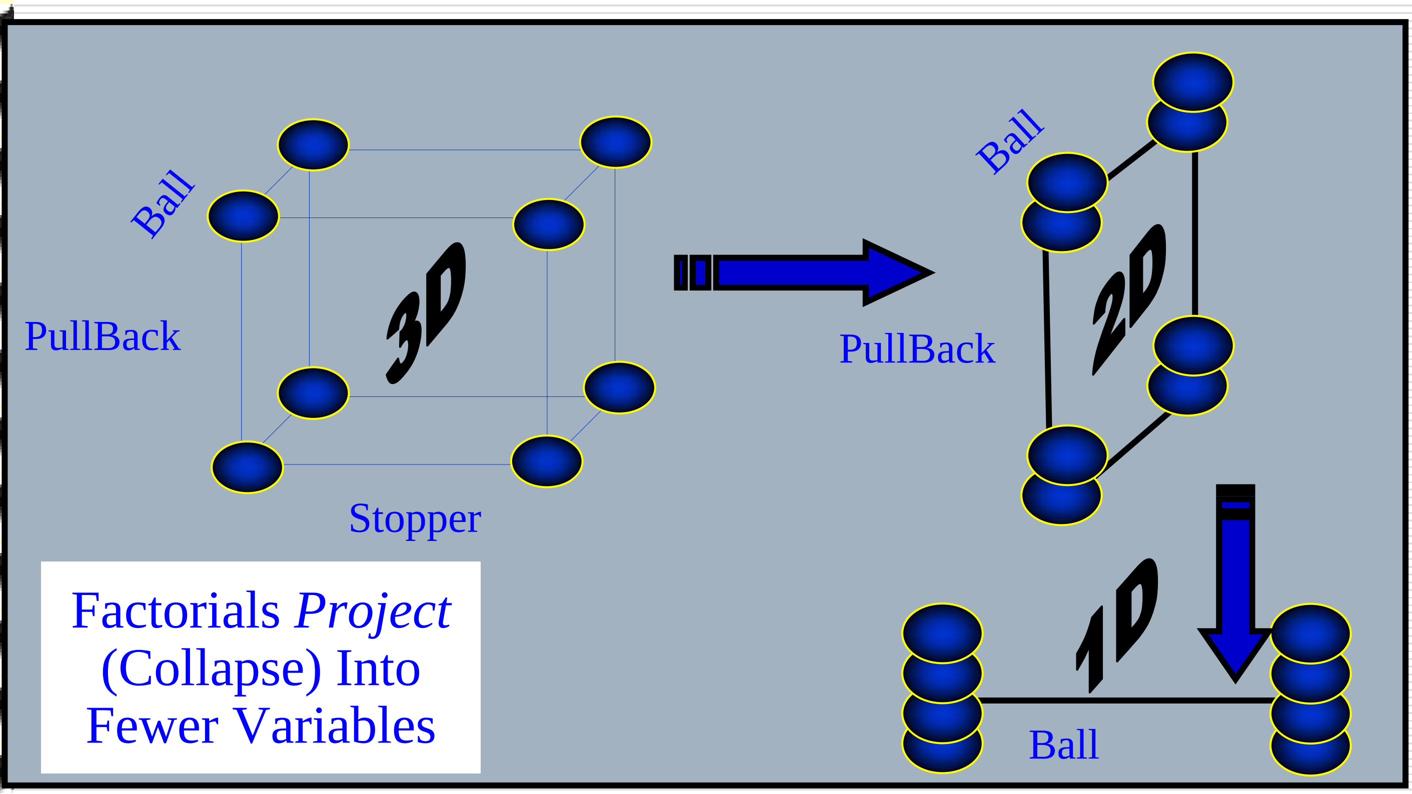
Formulation Sample -- Factorials (Solution)

Case	Ball	Tgt	PullBack
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	-1	1	1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1

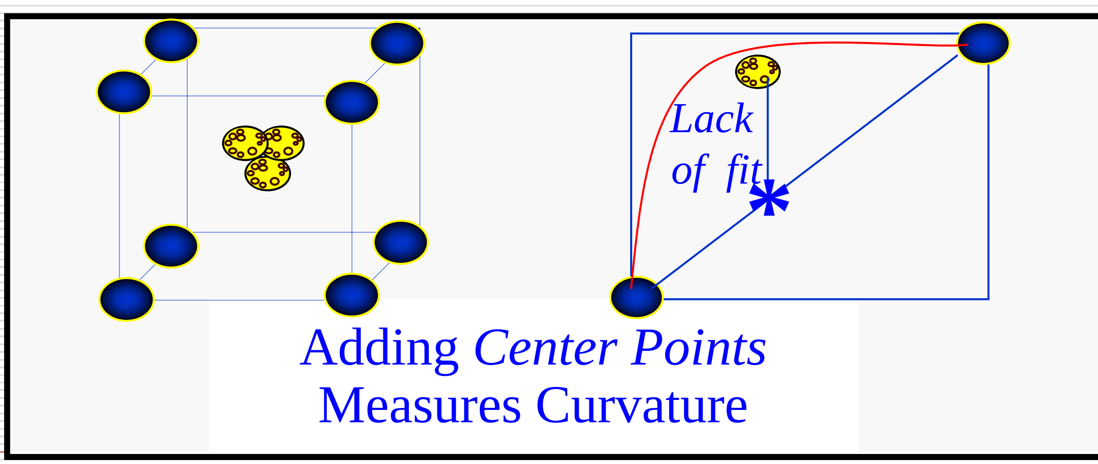


- Note the geometric shape you built
- All variables are *orthogonal* (right angles)
- General algorithm to build -- vary A for half of runs, B for half of A and C for half of B …

Some Properties of Factorial Designs



- Two level factorials are powerful for exploring
- Form the basis for many more elegant designs
- Can be augmented to fit more complicated models



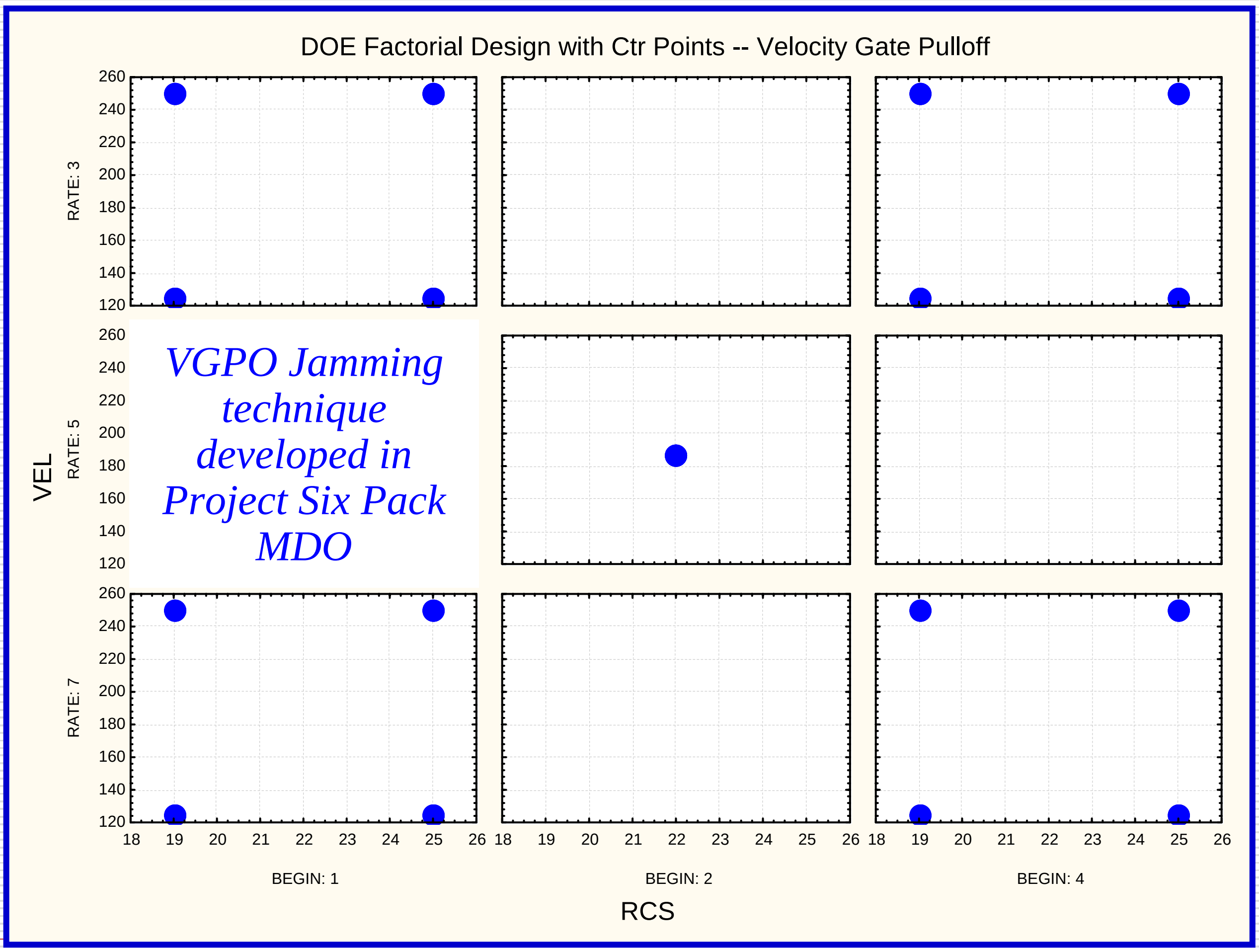
Flexibility: DOE Designs can be Simple or Complex

	Dive Angle		
Altitude	Level	-20	+30
Low	SLD	LAT	LOFT
High	HARB	Dive HARB	NA

Unguided bomb matrix from F-15E Suite 4E+ FDE

- Lack of understanding of this point leads to statements like :

"DOE is OK for ___ but cannot do ___"



Larger Factorial: 4 Factor Design

	Design				Interactions										
Case	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	**-1**	**-1**	**-1**	**-1**	1	1	1	1	1	1	-1	-1	-1	-1	1
2	**-1**	**-1**	**-1**	**1**	1	1	1	-1	-1	-1	-1	1	1	1	-1
3	**-1**	**-1**	**1**	**-1**	1	-1	-1	1	1	-1	1	-1	1	1	-1
4	**-1**	**-1**	**1**	**1**	1	-1	-1	-1	-1	1	1	1	-1	-1	1
5	**-1**	**1**	**-1**	**-1**	-1	1	-1	1	-1	1	1	1	-1	1	-1
6	**-1**	**1**	**-1**	**1**	-1	1	-1	-1	1	-1	1	-1	1	-1	1
7	**-1**	**1**	**1**	**-1**	-1	-1	1	1	-1	-1	-1	1	1	-1	1
8	**-1**	**1**	**1**	**1**	-1	-1	1	-1	1	1	-1	-1	-1	1	-1
9	**1**	**-1**	**-1**	**-1**	-1	-1	1	-1	1	1	1	1	1	-1	-1
10	**1**	**-1**	**-1**	**1**	-1	-1	1	1	-1	-1	1	-1	-1	1	1
11	**1**	**-1**	**1**	**-1**	-1	1	-1	-1	1	-1	-1	1	-1	1	1
12	**1**	**-1**	**1**	**1**	-1	1	-1	1	-1	1	-1	-1	1	-1	-1
13	**1**	**1**	**-1**	**-1**	1	-1	-1	-1	-1	1	-1	-1	1	1	1
14	**1**	**1**	**-1**	**1**	1	-1	-1	1	1	-1	-1	1	-1	-1	-1
15	**1**	**1**	**1**	**-1**	1	1	1	-1	-1	-1	1	-1	-1	-1	-1
16	**1**	**1**	**1**	**1**	1	1	1	1	1	1	1	1	1	1	1

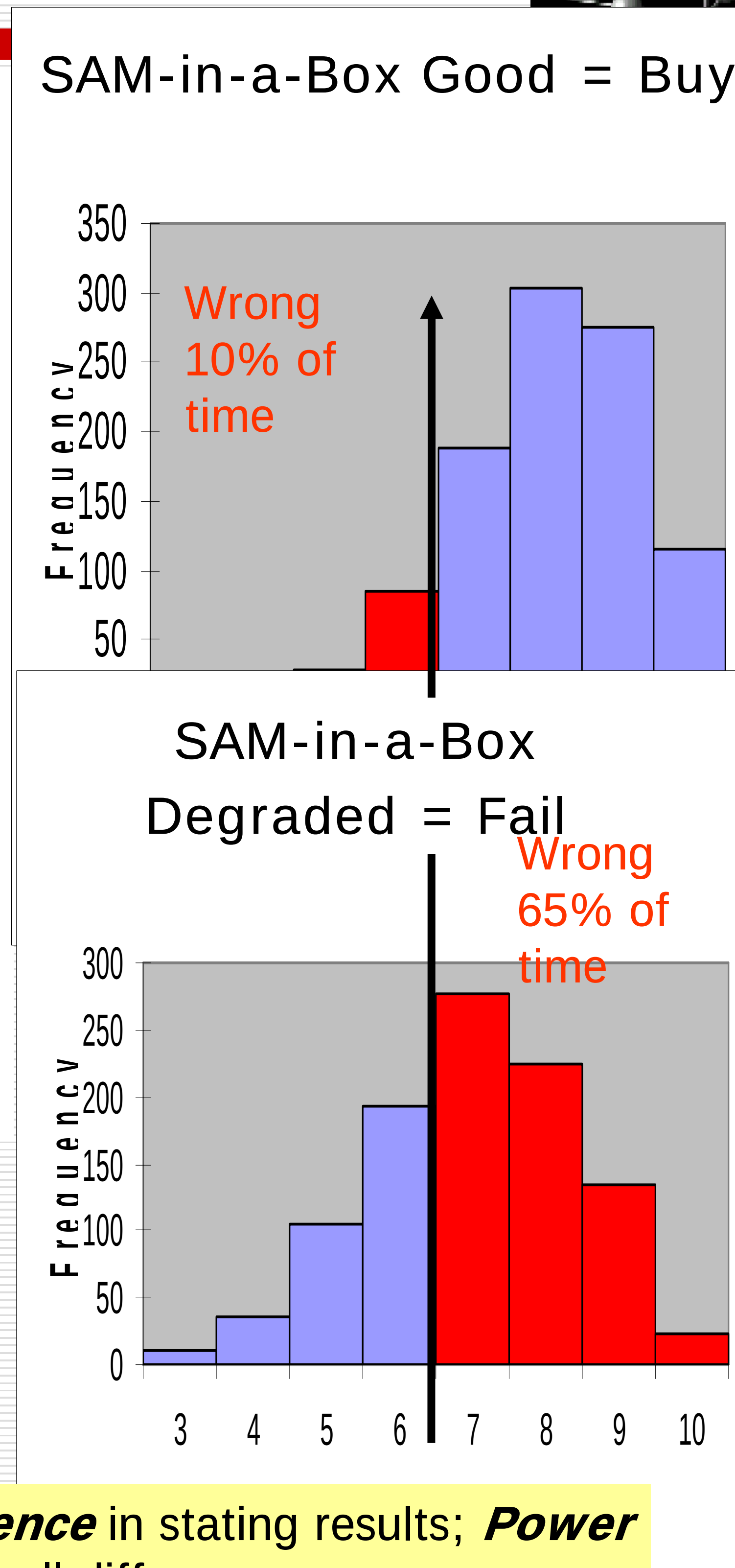
Effects and Interactions in *Standard Order*

Case	A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
1	**-1**	**-1**	**1**	**-1**	1	1	-1	-1	1	1	-1	1	-1	-1	1
2	**-1**	**-1**	**1**	**-1**	1	1	-1	1	-1	-1	1	-1	1	1	-1
3	**-1**	**-**												1	-1
4	**-1**	**-**												-1	1
5	**-1**	**1**												1	-1
6	**-1**	**1**												-1	1
7	**-1**	**1**	**-1**	**1**	-1	1	-1	-1	1	-1	1	-1	1	-1	1
8	**-1**	**1**	**-1**	**1**	-1	1	-1	1	-1	1	-1	1	-1	1	-1
9	**1**	**-1**	**-1**	**-1**	-1	1	1	-1	-1	1	1	1	1	-1	-1
10	**1**	**-1**	**-1**	**-1**	-1	1	1	1	1	-1	-1	-1	-1	1	1
11	**1**	**-1**	**-1**	**1**	1	-1	-1	-1	-1	1	1	-1	-1	1	1
12	**1**	**-1**	**-1**	**1**	1	-1	-1	1	1	-1	-1	1	1	-1	-1
13	**1**	**1**	**1**	**-1**	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
14	**1**	**1**	**1**	**-1**	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
15	**1**	**1**	**1**	**1**	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
16	**1**	**1**	**1**	**1**	1	1	1	1	1	1	1	1	1	1	1

- Standard Order (adding factors+interactions) helps keep terms straight:
  Y=M+A+B+AB+C+AC+BC+ABC+....

Recall -- we seek to balance our chance of errors

- Putting them together, we see we can trade one error for the other (α for β)
- We can also increase sample size to decrease our risks in testing
- These statements are not opinion – they are <u>mathematical fact</u> and an inescapable challenge in testing



Getting it right: ***Confidence*** in stating results; ***Power*** to find small differences

Replicating a Design – Simple rules for sample size

Percent confidence that identified effect exists -- (1-α)	α	Percent chance of finding true effects (1-β)	β	Total Unique Cases in Factorial Matrix: 2	4	8	12	16
95%	5%	40%	60%	3	2	1		
		75%	25%	5	3	2	1	1
		90%	10%	7	4	2	2	2
		95%	5%	8	5	3	2	2
		99%	1%	11	6	3	3	2
99%	1%	15%	85%	3	2	1		
		50%	50%	5	3	2		1
		70%	30%	7	4	2	2	2
		90%	10%	8	5	3	2	2
		96%	4%	11	6	3	3	2

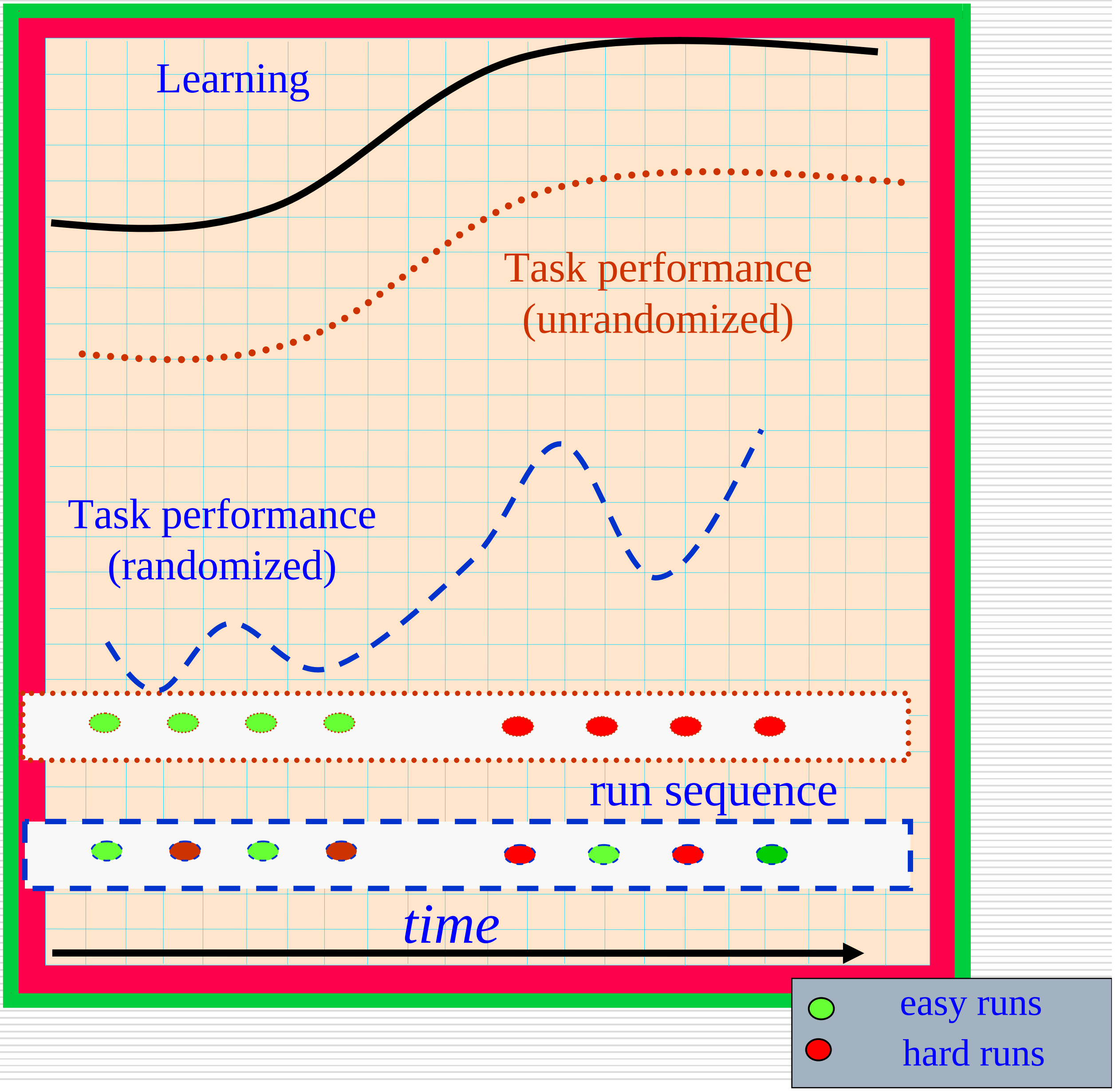
Samples Required Per Case

Illustration: 3x2x2 matrix gives 12 cases. With $\alpha = \beta = .05$, n (samples required per case) = ______

WARNING: *Only* applicable to Factorial Designs analyzed with these methods!`

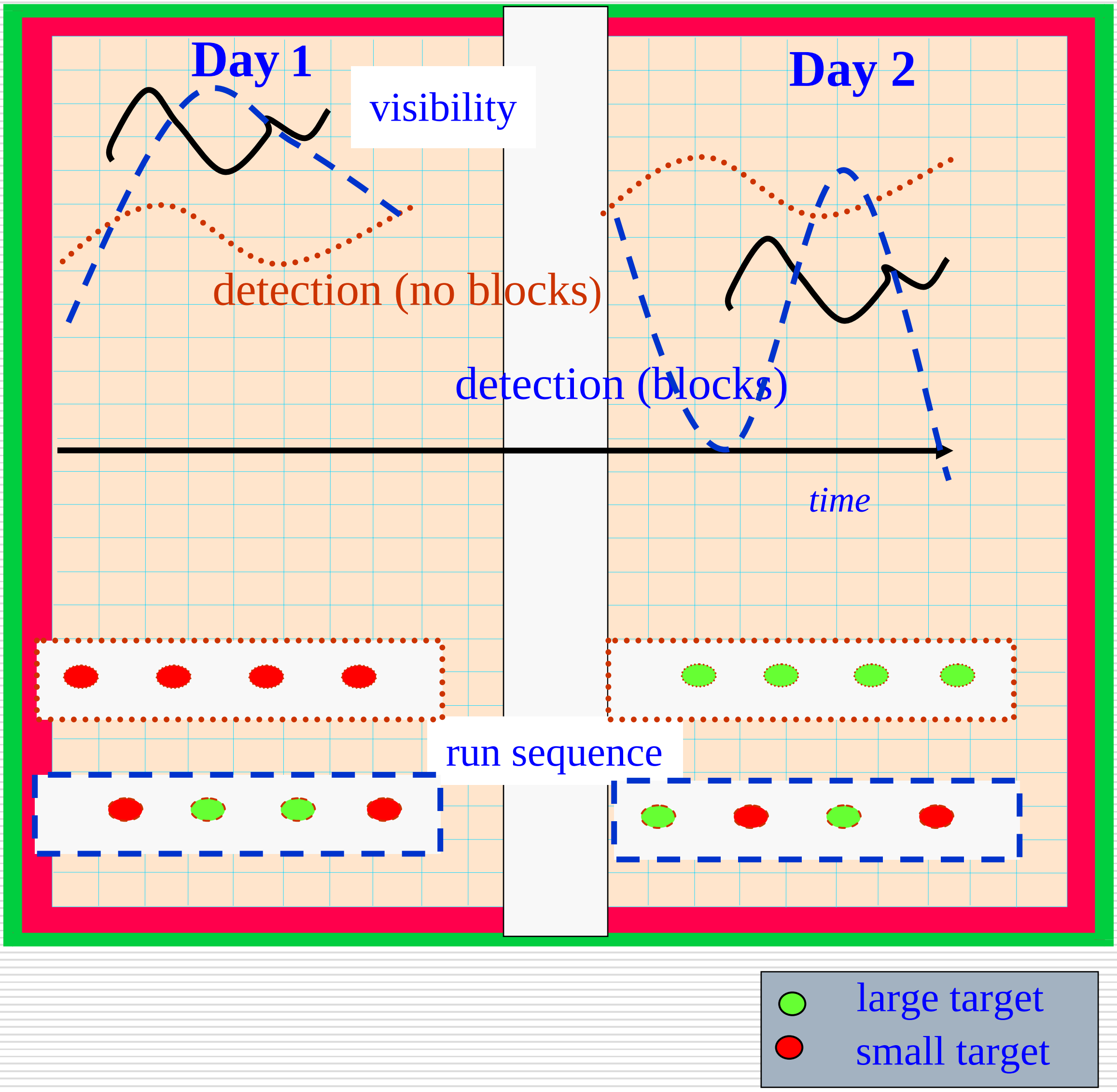
Source: Appendix M-2, *Understanding Industrial Designed Experiments*, Schmidt and Launsby, Air Academy Associates, 2000

Stage III PRODUCE Randomized Observations



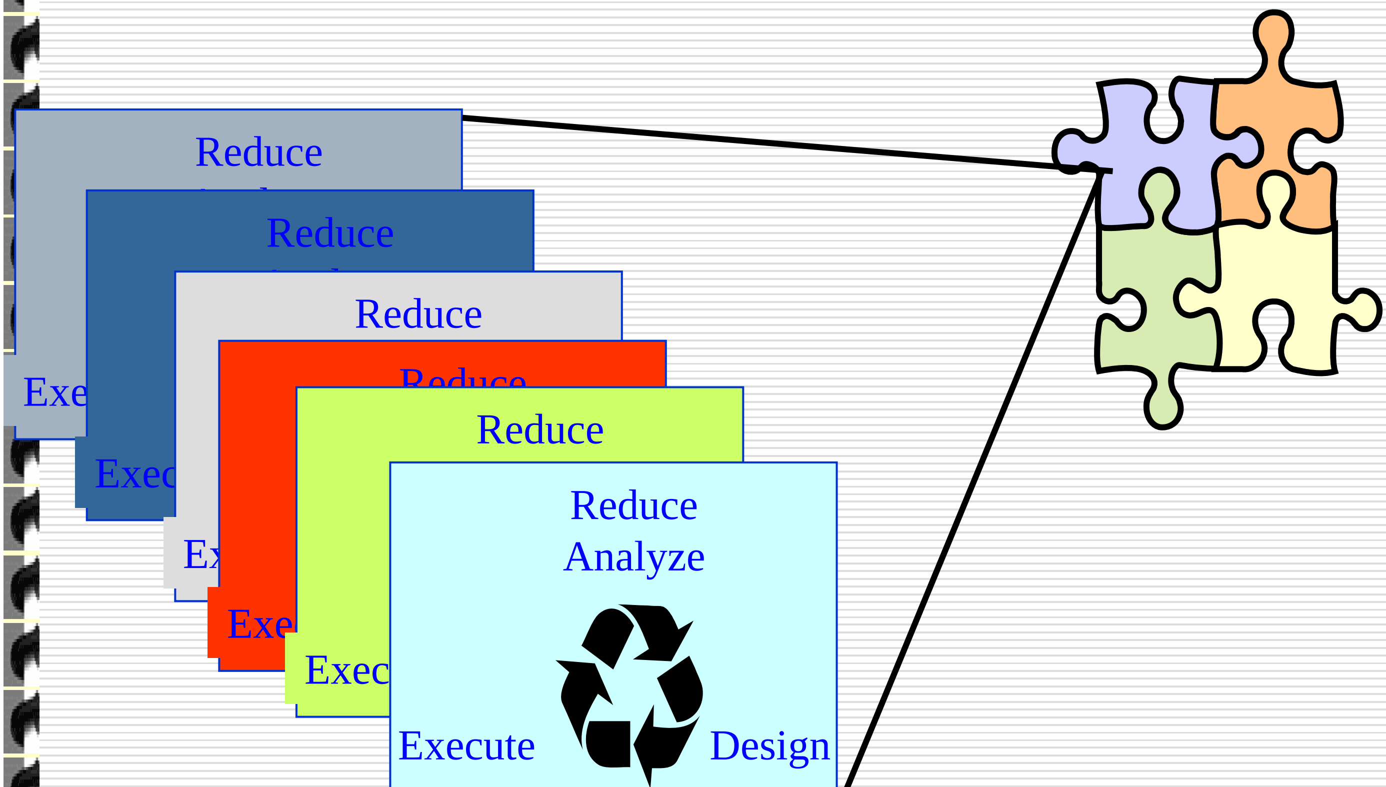
- **Randomizing** runs protects from unknown background changes **within** an experimental period (due to Fisher)

PRODUCE observations in Blocks



- **Blocking** <u>designs</u> protects from unknown background changes **between** experimental periods (also due to Fisher)

Simple rules for randomization and blocking



- Randomize each set of runs within a single test block, phase or mission
- Run one replicate of the design for each block or mission (produces a "block")
- Or -- run a similar set of ***calibration*** runs for each mission -- 4 to 8 trials and compare among blocks to detect different performance

Full Factorial Designs

1. Review your cause-and-effect diagram for the factors affecting variability of the statapult. Based on the cause-and-effect, develop standard operating procedures to control variability.

2. For this experiment, we will test 3 factors of interest: 1) stop position, 2) launch angle, 3) tension, and 4) ball type. Each factor will have two levels according to the following table. Be sure to hold all the other factors constant, if possible.

Factor	Low	High
Stop	2	4
Launch	160	180
Tension	2	4
Ball	Type 1	Type 2

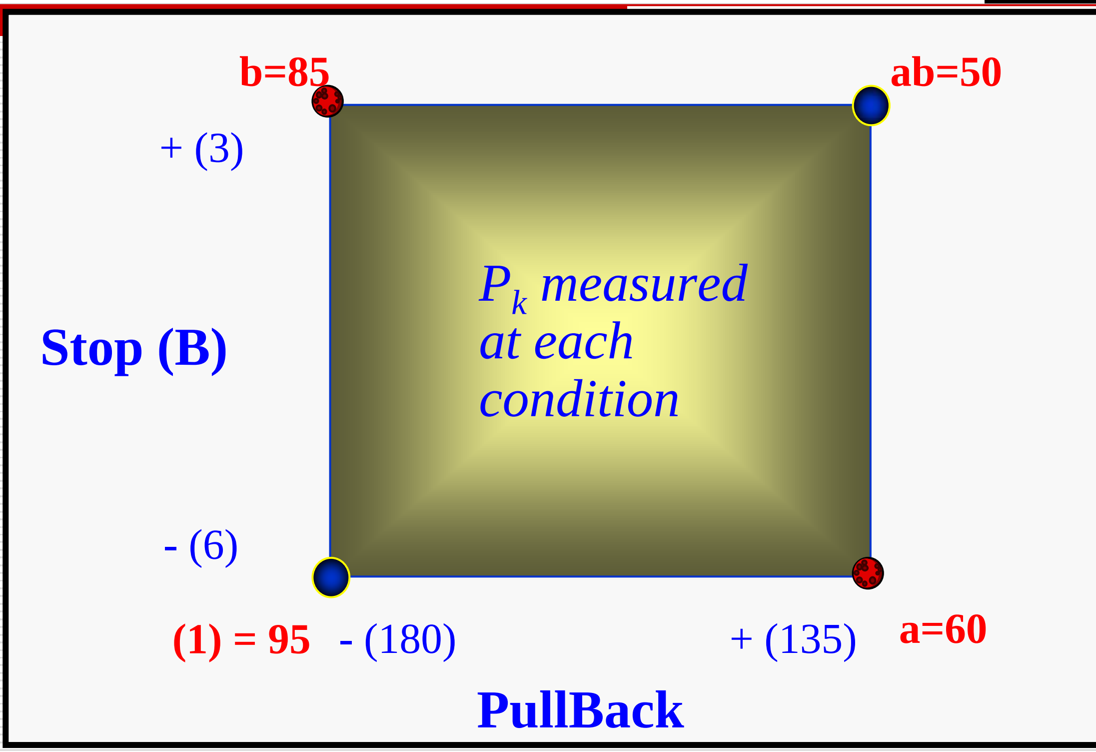
3. Use a 2^4 factorial design approach with 1 replicate. Develop a design matrix using the software to determine the random order of the runs.

4. Perform the experiment and collect the data. Make observations concerning held constant and nuisance factors during the experiment for future use.

5. Analyze your data. Determine the significant effects. Develop a predictive model of launch distance. Comment on your findings below.

6. Run confirmation runs using distances provided by the instructor.

Stage IV – PONDER
Results (Simple Analysis)



- How do Pullback & Stop affect Rng?
- Effect (Pullback) =
- Effect (Stop) =
- Effect (Pullback X Stop) =

***Designed Experiment** (n). Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).*

	Factorial Effect		
Case	**A**	**B**	**AB**
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

$$\hat{y} = \bar{\bar{y}} + \Delta_A + \Delta_B + \Delta_{AB}$$

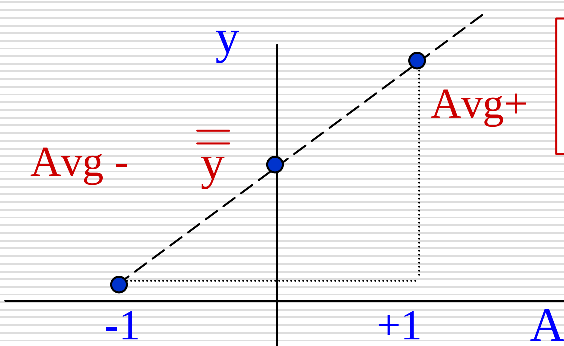
Why this math model works from Algebra II

$$\hat{y} = \overline{\overline{y}} + \frac{\Delta_A}{2}A + \frac{\Delta_B}{2}B + \frac{\Delta_{AB}}{2}AB$$

Where does equation come from?

1. Orthogonal design
2. Orthogonal coding
3. 2 levels for each factor.

Simple y=mx+b from Algebra II

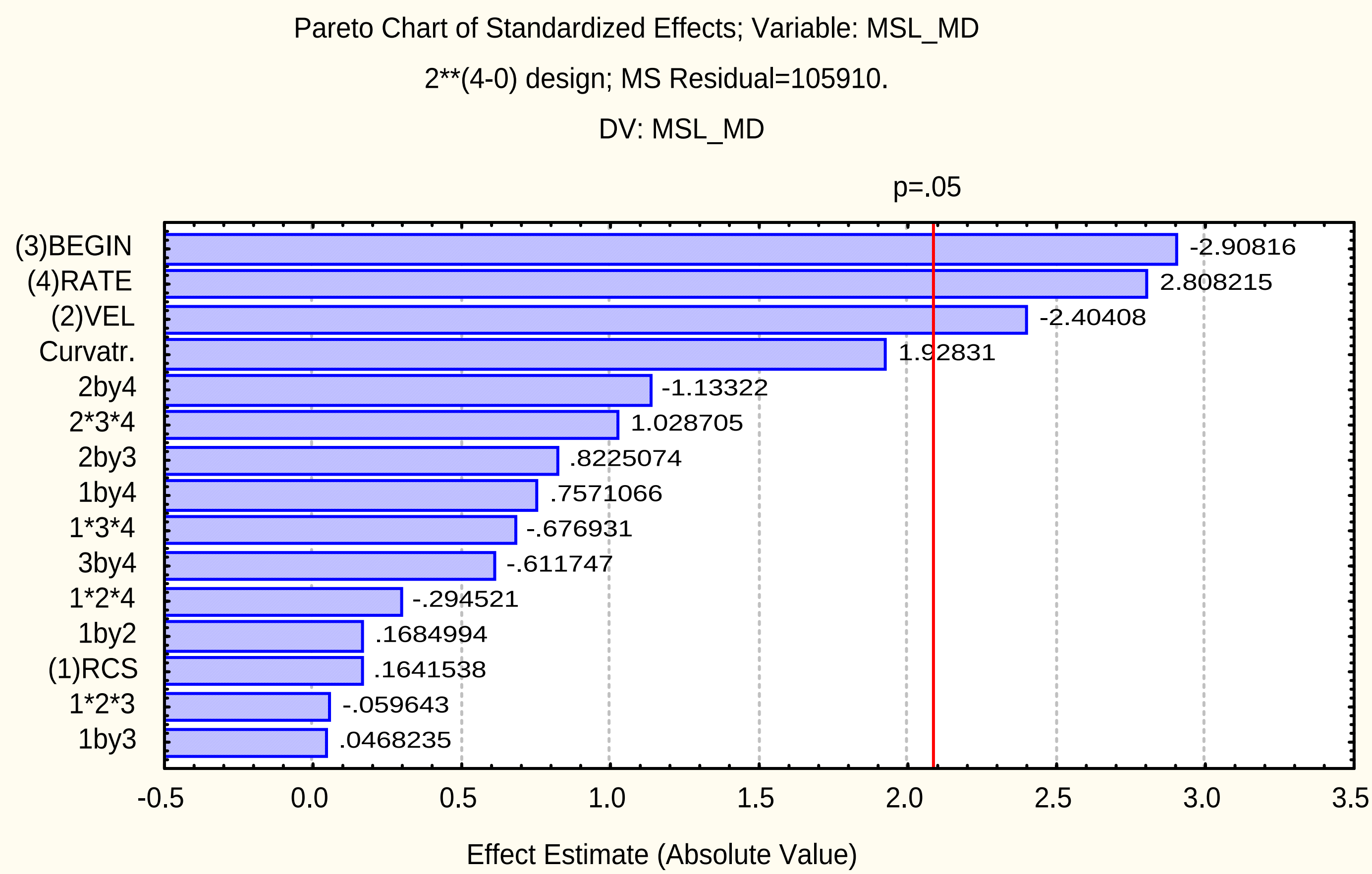


$$\overline{\overline{y}} = \left(\frac{Avg_{+} - Avg_{-}}{2}\right)A$$

Prediction equation for one factor (A)

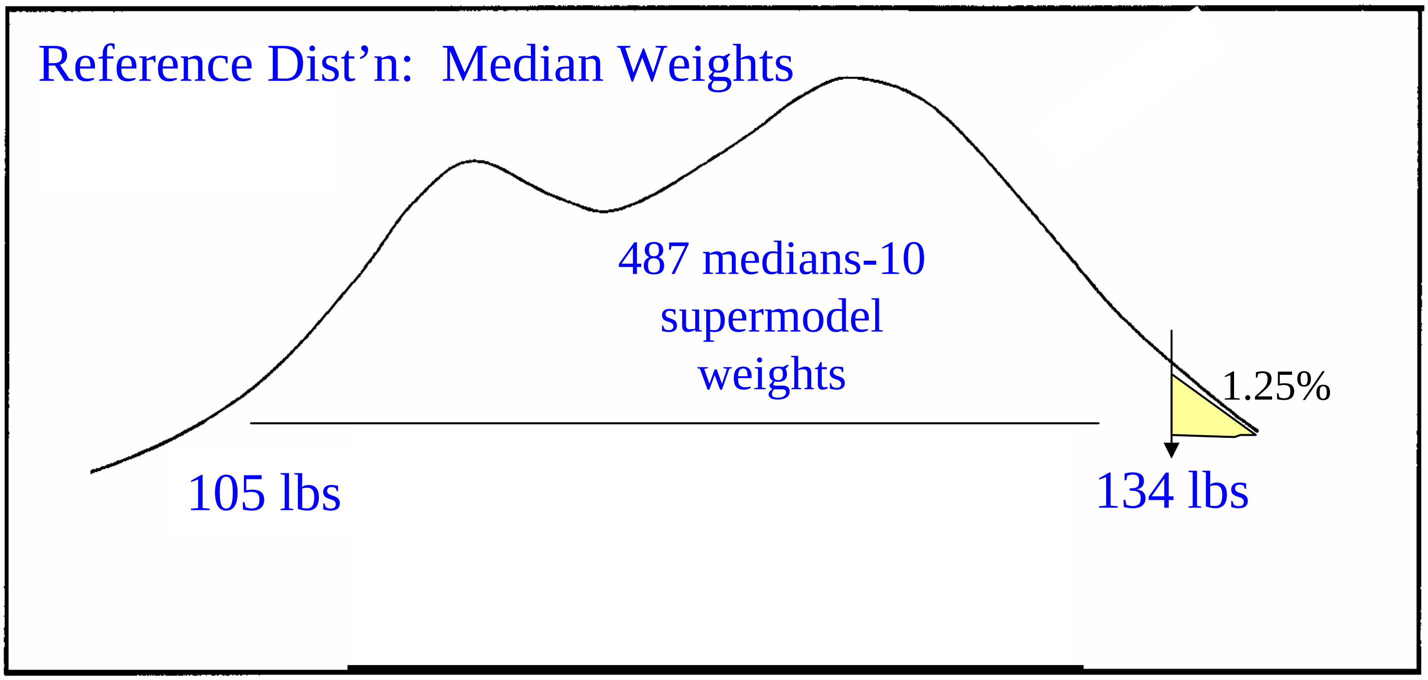
We can add other factors because our designs are orthogonal (predictors are independent of each other)

Statistics needed with more complex problems … Where is noise floor?

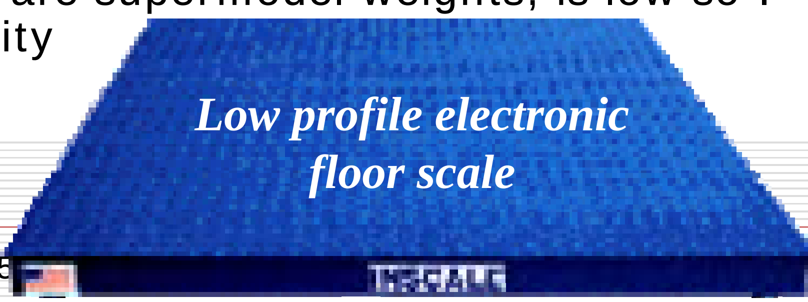


- Actual Statistica Pareto chart of recent test of an ECM technique development test. Four factors, each at two levels gives 16 combinations and 15 possible main effects and interactions.
- Red line drawn with chance of a false positive (α error) equal to 5%.

Concept of an External Reference Distribution



- Suppose I am a covert supermodel weigher. I travel the world from Milan to New York to clandestinely weigh supermodels in groups of 10
- Someone offers me a data set claiming it's supermodel weight data. X_{50} = 134 lbs, n = 10
- Examining my reference collection, I observe that only 1.25% of my 487 samples equal or exceed 134 lbs
- Accordingly, I reject this set with a 1.25% change of being mistaken
- I find likelihood (α) that I see a median of 134 lbs given these are supermodel weights, is low so I reject equality



There are 16 terms in our model – are all active?

	Design				Interactions										
Case	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	**-1**	**-1**	**-1**	**-1**	1	1	1	1	1	1	-1	-1	-1	-1	1
2	**-1**	**-1**	**-1**	**1**	1	1	1	-1	-1	-1	-1	1	1	1	-1
3	**-1**	**-1**	**1**	**-1**	1	-1	-1	1	1	-1	1	-1	1	1	-1
4	**-1**	**-1**	**1**	**1**	1	-1	-1	-1	-1	1	1	1	-1	-1	1
5	**-1**	**1**	**-1**	**-1**	-1	1	-1	1	-1	1	1	1	-1	1	-1
6	**-1**	**1**	**-1**	**1**	-1	1	-1	-1	1	-1	1	-1	1	-1	1
7	**-1**	**1**	**1**	**-1**	-1	-1	1	1	-1	-1	-1	1	1	-1	1
8	**-1**	**1**	**1**	**1**	-1	-1	1	-1	1	1	-1	-1	-1	1	-1
9	**1**	**-1**	**-1**	**-1**	-1	-1	1	-1	1	1	1	1	1	-1	-1
10	**1**	**-1**	**-1**	**1**	-1	-1	1	1	-1	-1	1	-1	-1	1	1
11	**1**	**-1**	**1**	**-1**	-1	1	-1	-1	1	-1	-1	1	-1	1	1
12	**1**	**-1**	**1**	**1**	-1	1	-1	1	-1	1	-1	-1	1	-1	-1
13	**1**	**1**	**-1**	**-1**	1	-1	-1	-1	-1	1	-1	-1	1	1	1
14	**1**	**1**	**-1**	**1**	1	-1	-1	1	1	-1	-1	1	-1	-1	-1
15	**1**	**1**	**1**	**-1**	1	1	1	-1	-1	-1	1	-1	-1	-1	-1
16	**1**	**1**	**1**	**1**	1	1	1	1	1	1	1	1	1	1	1

Effects and Interactions in Standard Order

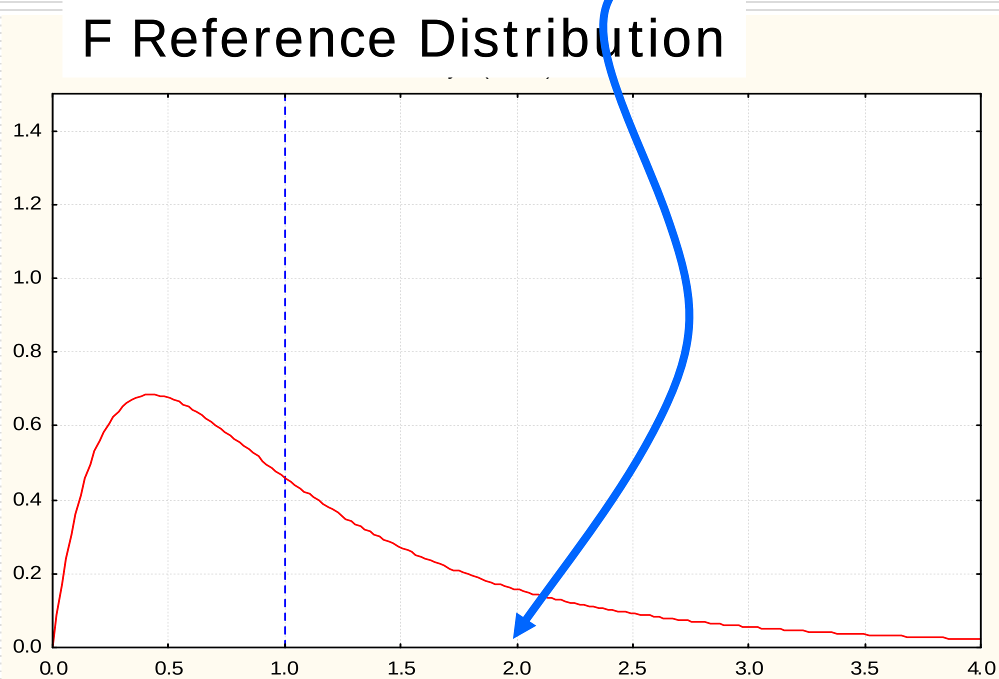
Case	A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
1	**-1**	**-1**	**1**	**-1**	1	1	-1	-1	1	1	-1	1	-1	-1	1
2	**-1**	**-1**	**1**	**-1**	1	1	-1	1	-1	-1	1	-1	1	1	-1
3	**-1**	**-1**												1	-1
4	**-1**	**-1**												-1	1
5	**-1**	**1**												1	-1
6	**-1**	**1**												-1	1
7	**-1**	**1**	**-1**	**1**	-1	1	-1	-1	1	-1	1	-1	1	-1	1
8	**-1**	**1**	**-1**	**1**	-1	1	-1	1	-1	1	-1	1	-1	1	-1
9	**1**	**-1**	**-1**	**-1**	-1	1	1	-1	-1	1	1	1	1	-1	-1
10	**1**	**-1**	**-1**	**-1**	-1	1	1	1	1	-1	-1	-1	-1	1	1
11	**1**	**-1**	**-1**	**1**	1	-1	-1	-1	-1	1	1	-1	-1	1	1
12	**1**	**-1**	**-1**	**1**	1	-1	-1	1	1	-1	-1	1	1	-1	-1
13	**1**	**1**	**1**	**-1**	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
14	**1**	**1**	**1**	**-1**	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
15	**1**	**1**	**1**	**1**	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
16	**1**	**1**	**1**	**1**	1	1	1	1	1	1	1	1	1	1	1

- Standard Order (adding factors+interactions) helps keep terms straight:
  Y=M+A+B+AB+C+AC+BC+ABC+....

ANOVA uses the F (for Fisher) Reference Distribution

Analysis of variance table

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	37008.25	4	9252.063	145.1304	< 0.0001
B	20449	1	20449	320.7686	< 0.0001
C	7482.25	1	7482.25	117.3686	< 0.0001
E	4147.2	1	4147.2	65.05412	< 0.0001
BE	1156	1	1156	18.13333	0.0013
Noise	701.25	11	63.75		
Cor Total	37709.5	15			



- In the Null world, F ~ 1.0
- Larger values of F lead to concluding differences exist

Break … we're limited to 8 shots/block – which 8 to run?

Std	A:Proj	B:Pullback	C:Stopper	D:FrFulcrum
1	Orange	150	2	1
2	Orange	150	2	3
3	Orange	150	4	1
4	Orange	150	4	3
5	Orange	180	2	1
6	Orange	180	2	
7	Orange	180	4	
8	Orange	180	4	3
9	Red	150	2	1
10	Red	150	2	3
11	Red	150	4	1
12	Red	150	4	3
13	Red	180	2	1
14	Red			
15	Red			
16	Red			

A full factorial 2^4 in 16 runs

Std	A:Proj	B:Pullback	C:Stopper	D:FrFulcrum
2	Orange	150	2	3
3	Orange	150	4	1
5	Orange	180	2	1
8	Orange	180	4	3
9	Red	150	2	1
12	Red	150	4	3
14	Red	180	2	3
15	Red	180	4	1
1	Orange	150	2	1
4	Orange	150	4	3
6	Orange	180	2	3
7	Orange	180	4	1
10	Red	150	2	3
11	Red	150	4	1
13	Red	180	2	1
16	Red	180	4	3

A full factorial 2^4 in 2 blocks of 8 runs each

Blocks – giving up some info to measure mission effect

With four factors – split 16 runs into two missions

- Step 1. Build a full factorial (16 runs) in all 4 vars
- Step 2. Construct high order interaction column (ABCD)
- Step 4. Confound (alias) ABCD with missions (blocks)

	Var>>	A	B	C	D	ABCD
Case	Block	Proj-ectile	Pull-Back	Stop-per	Fulcrum	
12	1	1	1	-1	1	-1
14	1	1	-1	1	1	-1
15	1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1
3	1	-1	1	-1	-1	-1
5	1	-1	-1	1	-1	-1
8	1	1	1	1	-1	-1
9	1	-1	-1	-1	1	-1
1	2	-1	-1	-1	-1	1
10	2	1	-1	-1	1	1
11	2	-1	1	-1	1	1
13	2	-1	-1	1	1	1
16	2	1	1	1	1	1
4	2	1	1	-1	-1	1
6	2	1	-1	1	-1	1
7	2	-1	1	1	-1	1

What can we learn from the second Block? - fraction of a full factorial

Given the number of factors k=4

Full fraction has 2^4=16 runs;

Half fraction has $2^{4-1} = 2^3 = 8$ runs

Seldom do this small half fraction since two-ways are confounded with each other

- Step 1. Build a full factorial (8 runs) in first 3 vars
- Step 2. Alias (perfectly confound) fourth factor with highest interaction (three way) in Step 1.
- Step 3. Determine Aliasing (confounding pattern) for all effects.

	Var>>	A	B	C	D=ABC
Case	Block	Proj	Pull	Stop	Fulc
1	2	-1	-1	-1	-1
13	2	-1	-1	1	1
11	2	-1	1	-1	1
7	2	-1	1	1	-1
10	2	1	-1	-1	1
6	2	1	-1	1	-1
4	2	1	1	-1	-1
[illegible]	2	1	[illegible]	[illegible]	1

Rules for Confounding Patterns:

I*A=A

$A*A=A^2=I$

Using this group Algebra--

D=ABC

I=ABCD

A=BCD

BC=AD ...

Five Variable Half Fraction

Case	A	B	C	D	E=ABCD
1	-1	-1	-1	-1	1
2	-1	-1	-1	1	-1
3	-1	-1	1	-1	-1
4	-1	-1	1	1	1
5	-1	1	-1	-1	-1
6	-1	1	-1	1	1
7	-1	1	1	-1	1
8	-1	1	1	1	-1
9	1	-1	-1	-1	-1
10	1	-1	-1	1	1
11	1	-1	1	-1	1
12	1	-1	1	1	-1
13	1	1	-1	-1	1
14	1	1	-1	1	-1
15	1	1	1	-1	-1
16	1	1	1	1	1

Step 1. Build a full factorial (2^{k-1} runs) in first k-1 variables

Step 2. Alias (perfectly confound) kth factor with highest interaction (2^{k-1} - way) in Step 1.

Step 3. Determine Aliasing (confounding pattern) for all effects.

Using group Algebra--
E=ABCD
I=ABCDE
other aliases?

The *Least* You Should Recall…

- Four Steps to DOE Include:
  - Process – decompose with flow & fishbone
  - Plan – build a factorial/crossed design
  - Produce – buy insurance with randomization & blocking
  - Ponder – use averages and ANOVA to draw conclusions
- Do not neglect power of simple tools
- Factorials solve the deep & broad problem
- Unknown-unknowns lurk
- Simple analysis uses geometry & slopes
- More complex blocks lead to idea of fractional factorials

Session 4 Fractional Factorials and Advanced Topics

- Overview of DOE Designs
- Fractional Factorials
- Dangerous Designs
  - Plackett Burman
  - Taguchi Orthogonal Arrays
  - Optimal Designs
- Variations on Factorials
  - Random effects, Nested, Split Plot, ANCOVA
- Mixed Models
- The 3-level model

Blocking (again)

Treatment Combo		Factorial Effect		
	I	A	B	C
(1)	+	-	-	-
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	-
c	+	-	-	+
ac	+	+	-	+
bc	+	-	+	+
abc	+	+	+	+

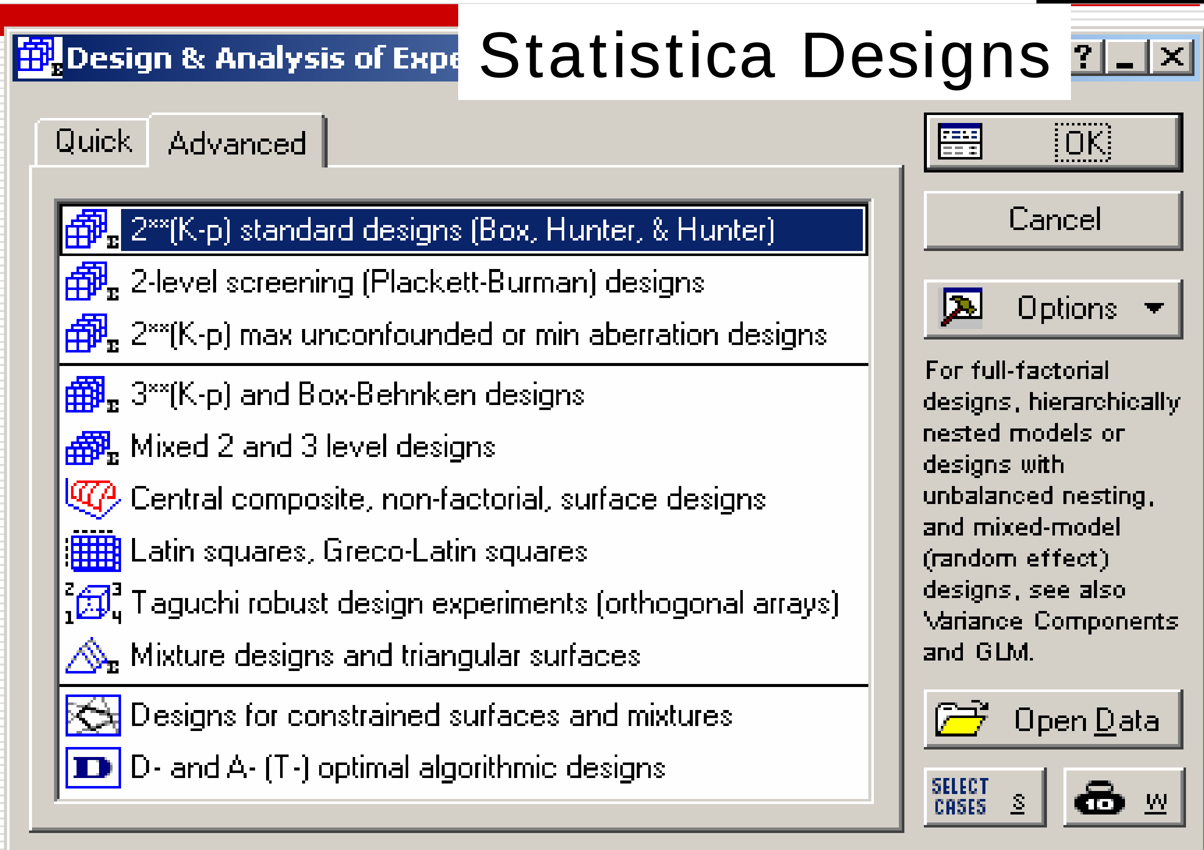
Blk_1	Blk_2	Blk_3
a	a	a
(1)	(1)	(1)
b	b	b
ab	ab	ab
c	c	c
ac	ac	ac
bc	bc	bc
abc	abc	abc

3 Reps of a 2^3 in 3 blocks

- Block - homogeneous experimental unit that
  - Potentially affects the response and
  - ***Restricts Randomization***
- Examples: missions, altitudes, pilots, aircraft, threat serial numbers, ….
- Simplest use of blocks is building up reps
- Example -- 2^3 design in a 12 pass mission

Let's Stroll Through Designs

Statistica Designs



- DE 6 Offers these designs
  - 2 Level factorials
  - Irregular Fractions
  - General factorials (including 3k designs)
  - D-Optimal designs
  - Plackett Burman screening designs
  - Taguchi Orthogonal Arrays
  - Plus:
    - (Mixture designs)
    - (Designs for constrained regions)

Blocking Replicates -- ANOVA Table

Source	df	SSquares	MSquare	F Statistic	P value
Blocks	bl-1=2	SSBlocks			
A (ECM)	1	SSA	MSA	FA	0.000
B (Altitude)	1	SSB	MSB	FB	0.000
C (Offset)	1	SSC	MSC	FC	0.000
AB	1	SSAB	MSAB	FAB	0.000
AC	1	SSAC	MSAC	FAC	0.000
BC	1	SSBC	MSBC	FBC	0.000
ABC	1	SSABC	MSABC	FABC	
Error	14	SSE	MSE		
Total	N-1=23				

- There are b-1 degrees of freedom associated with the blocks and b blocks.
- Note no block-treatment interaction
- Note no block F statistic

2^2 Factorial in 2 Blocks

T'ment Combo	I	A	B	AB
	Factorial Effect			
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

Blk_1	Blk_2
(1)	a
ab	b

A= ab+a-(1)-b
B=ab+b-(1)-a
AB=ab+(1)-a-b

Blocks=?

- Confound (or intentionally intertwine) AB with Blocks
- Note that A and B are unaffected by blocks -- there are one plus and minus in each block and the block effect will cancel out

2^3 in Two Blocks

	Factorial Effect							
Treatment Combo	I	A	B	AB	C	AC	BC	ABC
(1)	+	-	-	+	-	+	+	-
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
ab	+	+	+	+	-	-	-	-
c	+	-	-	+	+	-	-	+
ac	+	+	-	-	+	+	-	-
bc	+	-	+	-	+	-	+	-
abc	+	+	+	+	+	+	+	+

Blk_1	Blk_2
(1)	a
ab	b
ac	c
bc	abc

- Confound ABC with blocks
  - Method 1 – choose +/- signs in ABC contrast
  - Method 2 – let stats program do it
  - There are two other methods …

2^k in 4 Blocks

				Blocking Generators					
Design: 2**(5-0) design in 4 Blocks				1	123			Statistica Generators	
				2	345			ABC	CDE
Block	**A**	**B**	**C**	**D**	**E**	**ABC**	**CDE**	L1	L2
1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
1	1	0	1	1	0	0	0	0	0
1	0	1	1	1	0	0	0	0	0
1	1	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0	0
1	0	0	0	1	1	0	0	0	0
1	1	1	0	1	1	0	0	0	0
2	1	0	1	0	0	0	1	0	1
2	0	1	1	0	0	0	1	0	1
2	0	0	0	1	0	0	1	0	1
2	1	1	0	1	0	0	1	0	1
2	0	0	0	0	1	0	1	0	1
2	1	1	0	0	1	0	1	0	1
2	1	0	1	1	1	0	1	0	1
2	0	1	1	1	1	0	1	0	1
3	1	0	0	0	0	1	0	1	0
3	0	1	0	0	0	1	0	1	0
3	0	0	1	1	0	1	0	1	0
3	1	1	1	1	0	1	0	1	0
3	0	0	1	0	1	1	0	1	0
3	1	1	1	0	1	1	0	1	0
3	1	0	0	1	1	1	0	1	0
3	0	1	0	1	1	1	0	1	0
4	0	0	1	0	0	1	1	1	1
4	1	1	1	0	0	1	1	1	1
4	1	0	0	1	0	1	1	1	1
4	0	1	0	1	0	1	1	1	1
4	1	0	0	0	1	1	1	1	1
4	0	1	0	0	1	1	1	1	1
4	0	0	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1

- Generate using two generators
- *blocks* have three df -- where is the third confounded factor?
- $ABC*CDE=ABC^2DE=ABDE$

Intro to Fractional Factorials

	Design: 2**(4-0) design (seatwork 10.sta)					
			A	B	C	D
		Block	(Cont.)	(Cont.)	(Cont.)	(Cont.)
I	8	1	1	1	1	-1
	1	1	-1	-1	-1	-1
	10	1	1	-1	-1	1
	15	1	-1	1	1	1
II	4	2	1	1	-1	-1
	11	2	-1	1	-1	1
	5	2	-1	-1	1	-1
	14	2	1	-1	1	1
III	6	3	1	-1	1	-1
	3	3	-1	1	-1	-1
	12	3	1	1	-1	1
	13	3	-1	-1	1	1
IV	7	4	-1	1	1	-1
	2	4	1	-1	-1	-1
	9	4	-1	-1	-1	1
	16	4	1	1	1	1

Guiding Principles Make Fractions Attractive

- Sparcity of Effects: There are 63 df in 2^6 design -- less than 10% of effects and interactions are usually “significant”
- Projection Property: Most fractional factorials project into a (possibly) replicated full factorial in fewer variables
- Sequential Experimentation: It is better to let facts drive you from design to design rather than opinion

Half Fraction 2^3 in ~~Two~~ One Block

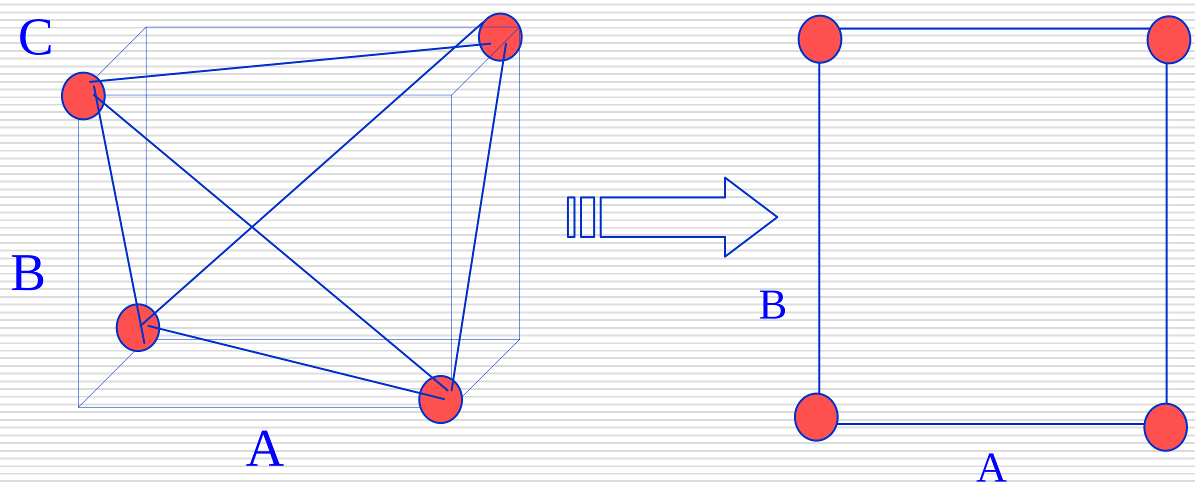
Design of Experiments
DOE

Treatment Combo	Factorial Effect							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	-	-	+	-	+	+	-
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
ab	+	+	+	+	-	-	-	-
c	+	-	-	+	+	-	-	+
ac	+	+	-	-	+	+	-	-
bc	+	-	+	-	+	-	+	-
abc	+	+	+	+	+	+	+	+

Blk_1	Blk_2
~~(1)~~	a
~~ab~~	b
~~ac~~	c
~~bc~~	abc

- Confound ABC with Fraction (same procedure as blocks)
  - Method 1 -- choose +/- signs in ABC contrast
  - Method 2 -- defining contrasts
- This is a 2^{3-1} design -- a half fraction of 2^3 -- four runs vs. 8
- ABC is our generator -- equal (in principal fraction) to I

Projection of Fractions



- This half fraction projects into full fraction of 2^2 design in any of three variables
- In general, a design of R, projects into possibly replicated factorial of R-1 variables
- Since max resolution of half fraction of 2^k is k, every 2^{k-1} fraction projects into full factorial in any k-1 factors, 2 reps in k-2, etc…
- **Not** generally true for higher fractions

Example Half Fraction

Temperature	Content	Treatment	Refiner	Length
-1	-1	-1	-1	1.71
1	-1	-1	1	1.86
-1	1	-1	1	1.79
1	1	-1	-1	1.67
-1	-1	1	1	1.81
1	-1	1	-1	1.25
-1	1	1	-1	1.46
1	1	1	1	0.85

- Suppose only one half fraction of problem 7-15 could be run 2^{4-1}
- Create design with DE6
- Run analysis using LEAP but Verify!
- Check results with original and half fraction of Replicate 2

One Quarter Fraction of 2^k Design (2^{k-2})

Label	A	B	C	D=AB	E=AC
(1)	-1	-1	-1	1	1
a	1	-1	-1	-1	-1
b	-1	1	-1	-1	1
ab	1	1	-1	1	-1
c	-1	-1	1	1	-1
ac	1	-1	1	-1	1
bc	-1	1	1	-1	-1
abc	1	1	1	1	1

Principal quarter fraction of 25 (2^{5-2})
R_{III}

I=ABD and I=ACE
Monty page 683

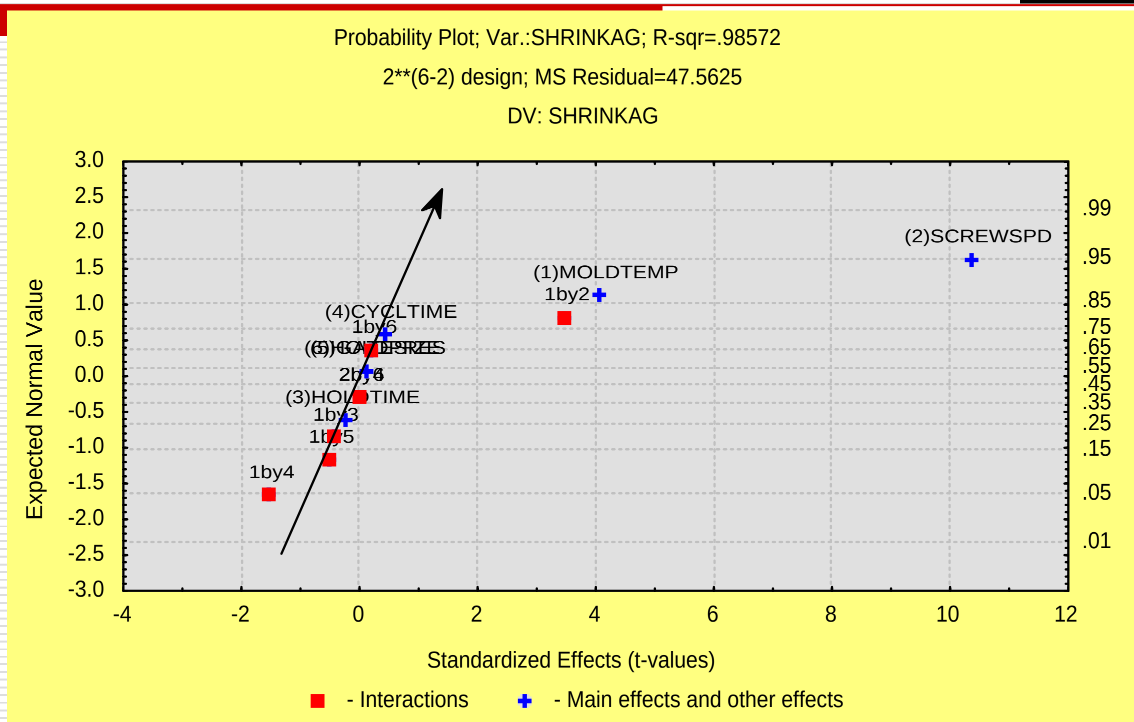
- Write FF Basic design in k-2 factors
- Confound (equate) k-1, k factors with 2 *appropriate* effects (e.g., P and Q)
- Compute settings for k, k-1 as effect products
- Generator is P, Q and PQ generalized interaction

Quarter Fraction Example

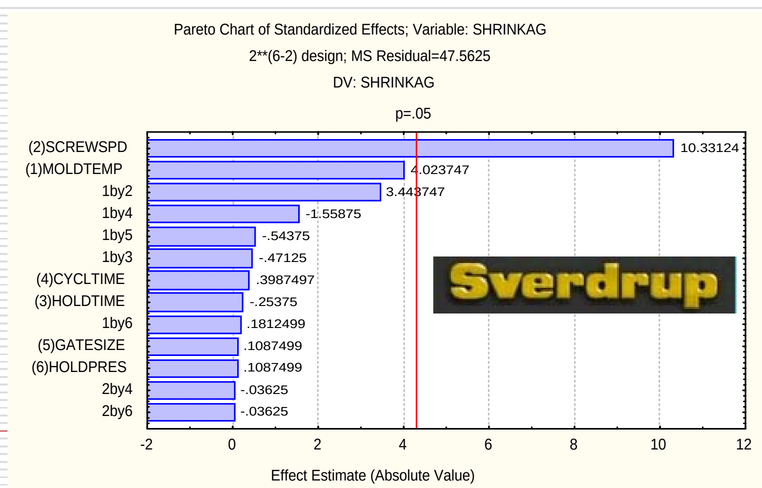
	Quarter Fraction of 26 design -- I=ABCE=BCDF=ADEF						
		--- Basic Design --------				Fractional Generators	
Case	Label	A	B	C	D	E=ABC	F=BCD
1	(1)	-1	-1	-1	-1	-1	-1
2	a	1	-1	-1	-1	1	-1
3	b	-1	1	-1	-1	1	1
4	ab	1	1	-1	-1	-1	1
5	c	-1	-1	1	-1	1	1
6	ac	1	-1	1	-1	-1	1
7	bc	-1	1	1	-1	-1	-1
8	abc	1	1	1	-1	1	-1
9	d	-1	-1	-1	1	-1	1
10	ad	1	-1	-1	1	1	1
11	bd	-1	1	-1	1	1	-1
12	abd	1	1	-1	1	-1	-1
13	cd	-1	-1	1	1	1	-1
14	acd	1	-1	1	1	-1	-1
15	bcd	-1	1	1	1	-1	1
16	abcd	1	1	1	1	1	1

- Montgomery example 8-4, a quarter fraction of 2^6 design: 2^{6-2}_{RIV}
- Generate in Statistica (careful of chosen generators)
- Objective -- minimize shrinkage after 24 hours
- Response -- percent shrinkage (transform??)
- Factors --
  - A-- Mold temperature B-- Screw speed, C-- hold time, D--Cycle time, E-- Gate size, F-- hold pressure

Analysis Results -- 3-ways assumed negligible



- A, B and AB stand out with AD as distant possibility
- Rule -- be flexible and humble but a little ruthless with p-values - you cannot consider everything!
- Get the Pareto results -- 80%



Res III Design Issues

	Res III Design								
CASE	A	B	C	D=+AB I=ABD	E=+AC I=ACE	F=+BC I=BCF	G=+ABC I=ABCG	Tr Combo	**Time**
1	-1	-1	-1	1	1	1	-1	def	85.5
2	1	-1	-1	-1	-1	1	1	afg	75.1
3	-1	1	-1	-1	1	-1	1	beg	93.2
4	1	1	-1	1	-1	-1	-1	abd	145.4
5	-1	-1	1	1	-1	-1	1	cdg	83.7
6	1	-1	1	-1	1	-1	-1	ace	77.6
7	-1	1	1	-1	-1	1	-1	bcf	95.0
8	1	1	1	1	1	1	1	abcdefg	141.8

	SS	**df**	**MS**
(1)A	850.781	1	850.781
(2)B	**2945.281**	**1**	**2945.281**
(3)C	0.151	1	0.151
(4)D_AB	**1667.531**	**1**	**1667.531**
(5)E_AC	0.151	1	0.151
(6)F_BC	0.781	1	0.781
(7)G_ABC	11.761	1	11.761
Error	0.000	0	
Total SS	5476.439	7	

	Confounding of Effec		
Factor	**Alias 1**	Alias 2	Alias 3
(1)A	2*4	3*5	6*7
(2)B	1*4	3*6	5*7
(3)C	1*5	2*6	4*7
(4)D_AB	1*2	3*7	5*6
(5)E_AC	1*3	2*7	4*6
(6)F_BC	1*7	2*3	4*5
(7)G_ABC	1*6	2*5	3*4

- Monty example 8-7, a human factors experiment 1/16th fraction of a 2^7 design.
- What can we do with the aliases of B & D?

Res III Foldover Designs

	Res III Design								
CASE	A	B	C	D=+AB I=ABD	E=+AC I=ACE	F=+BC I=BCF	G=+ABC I=ABCG	Tr Combo	**Time**
1	-1	-1	-1	1	1	1	-1	def	85.5
2	1	-1	-1	-1	-1	1	1	afg	75.1
3	-1	1	-1	-1	1	-1	1	beg	93.2
4	1	1	-1	1	-1	-1	-1	abd	145.4
5	-1	-1	1	1	-1	-1	1	cdg	83.7
6	1	-1	1	-1	1	-1	-1	ace	77.6
7	-1	1	1	-1	-1	1	-1	bcf	95.0
8	1	1	1	1	1	1	1	abcdefg	141.8

- Monty 8-7 -- Problems with aliases, projection
- Solution – Fold over by Reversing the signs in generators with *odd* number of letters in their word
- The *odd* letters alias mains with 2-ways
- This is a fold over design RIII => RIV

	Plus Complete Foldover Gives Res IV								
CASE	A	B	C	**D=-AB I=-ABD**	**E=-AC I=-ACE**	**F=-BC I=-BCF**	**G=+ABC I=+ABCG**	Tr Combo	**Time**
9	-1	-1	-1	-1	-1	-1	-1	(1)	71.9
10	1	-1	-1	1	1	-1	1	adeg	87.3
11	-1	1	-1	1	-1	1	1	bdfg	143.8
12	1	1	-1	-1	1	1	-1	abef	94.1
13	-1	-1	1	-1	1	1	1	cefg	73.4
14	1	-1	1	1	-1	1	-1	acdf	82.4
15	-1	1	1	1	1	-1	-1	bcde	136.7
16	1	1	1	-1	-1	-1	1	abcg	91.3

Results of the Fold Over

	SS	df	MS	F	p
(1)A	8.70	1	8.7	0.52	0.60
(2)B	**5791.21**	**1**	**5791.2**	**344.51**	**0.03**
(3)C	12.96	1	13.0	0.77	0.54
(4)D_AB	**3451.56**	**1**	**3451.6**	**205.33**	**0.04**
(5)E_AC	0.06	1	0.1	0.00	0.96
(6)F_BC	1.00	1	1.0	0.06	0.85
(7)G_ABC	0.06	1	0.1	0.00	0.96
1 by 2	1.00	1	1.0	0.06	0.85
1 by 3	0.64	1	0.6	0.04	0.88
1 by 4	0.42	1	0.4	0.03	0.90
1 by 5	9.30	1	9.3	0.55	0.59
1 by 6	26.01	1	26.0	1.55	0.43
1 by 7	5.06	1	5.1	0.30	0.68
2 by 4	**1466.89**	**1**	**1466.9**	**87.26**	**0.07**
Error	16.81	1	16.8		
Total SS	10791.70	15			

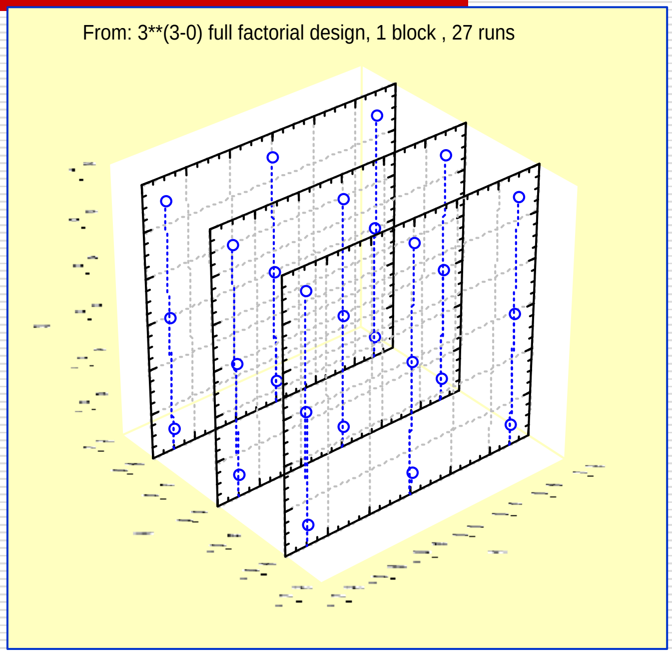
	Confounding	
Factor	**Alias 1**	Alias 2
(1)A		
(2)B		
(3)C		
(4)D_AB		
(5)E_AC		
(6)F_BC		
(7)G_ABC		
1 by 2	3*7	5*6
1 by 3	2*7	4*6
1 by 4	3*6	5*7
1 by 5	2*6	4*7
1 by 6	2*5	3*4
1 by 7	2*3	4*5
2 by 4	3*5	6*7

- Res IV design indicates it's B, D and BD interaction
- B and D de-aliased with two factor interactions

9 ~~8~~ Simple Rules for De-Aliasing

1. Run the other half fraction
2. Augment fraction to break selected chains
   - (Hidinger Conjecture – 2 runs break 2 chains
3. Assume away 3-way (4-Way) and higher interactions
4. Subject matter reasoning – this one, not that
5. Collapse across inactive var & project
6. Predict and confirm results
7. Foldover RIII to RIV (Complete foldover)
8. Foldover on a letter: RIII -> RV one letter
9. Ockham's razor – big main effects => guilty of big two ways

Three level designs have drawbacks

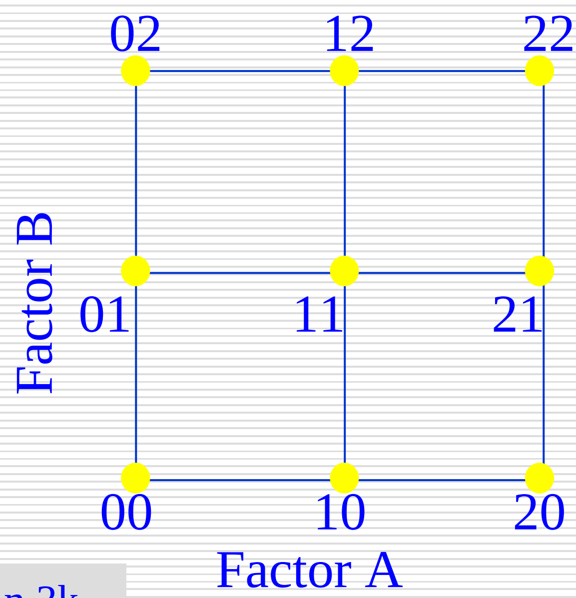


- **Sometimes used when 2nd order models or curvature are desired**
  - Response surface designs (central composite designs) superior
  - 2 level factorial with center points ok for curvature
- Details
  - Each variable A, B, C,...k has three levels 0,1,2
  - Total of 3^k cases in design
  - Main effects each have 2 df vs. 1 in 2^k designs
  - Can be used to estimate linear and quadratic elements

Factors at 3 Levels

3**(2-0) full factorial design,...

EXPERIM DESIGN	A	B
1	-1	-1
2	-1	0
3	-1	1
4	0	-1
5	0	0
6	0	1
7	1	-1
8	1	0



New notation: each treatment combo in 3k has k digits -- 0,1,2 denoting factor levels.

All at low denoted 000

All high denoted 222

A high, B med C low denoted 210

- New notation -- 012 or familiar -101 to denote levels
- For 3^2 design -- 9 runs with 8 df
- 2-way interactions (e.g. AB) have (a-1)*(b-1)=4 df
  - partitioned into LxL LxQ QxL and QxQ
  - Possible to partition in AB and AB^2 (I,J) also

Example 3^2 design

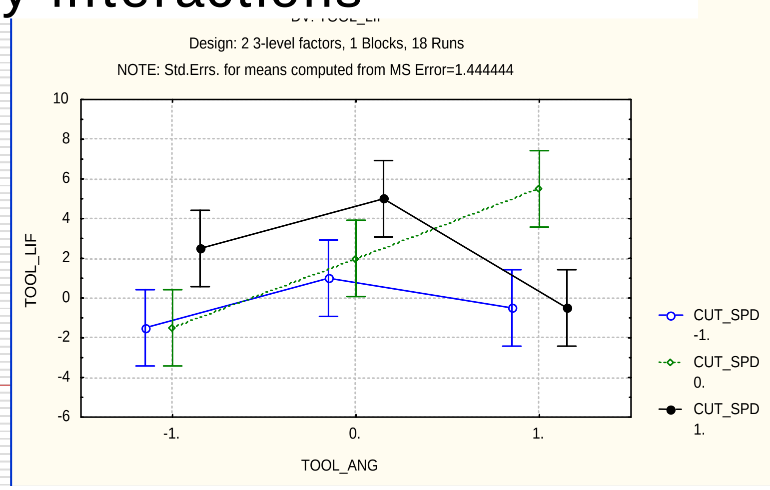
Effect	SS	df	MS	F	p
(1)TOOL_ANG L+Q	24.	[illegible]	12.17	8.42	0.009
(2)CUT_SPD L+Q	25.33	2	12.67	8.77	0.008
1*2	61.33	4	15.33	10.62	0.002
Error	13	9	1.44		
Total SS	124	17			

Usual ANOVA

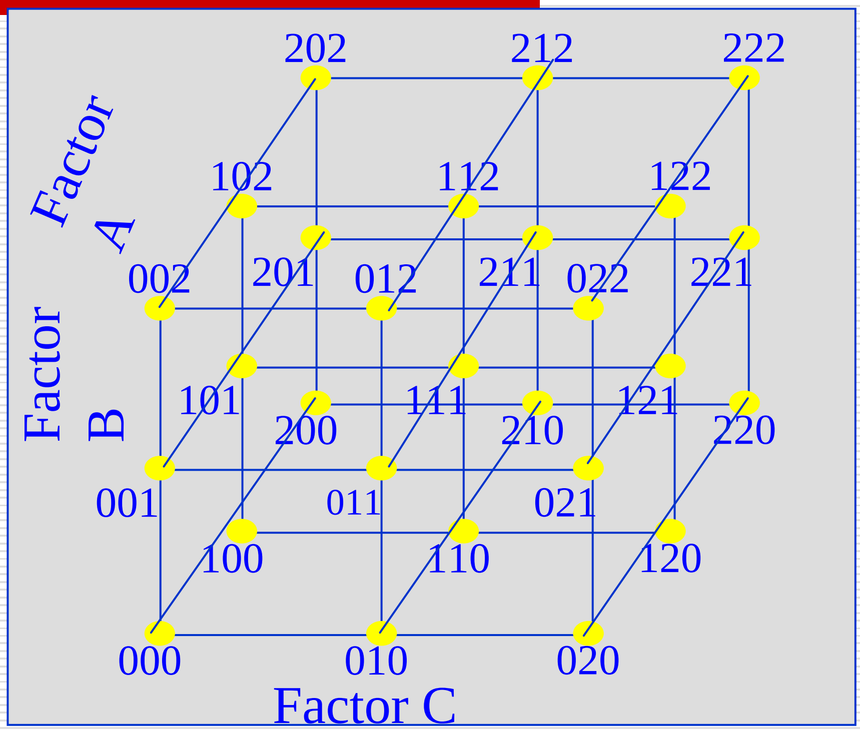
Effect	SS	df	MS	F	p
(1)TOOL_ANG(L)	[illegible]	[illegible]	8.3	5.8	0.040
TOOL_ANG(Q)	16.0	1	16.0	11.1	0.009
(2)CUT_SPD (L)	21.3	1	21.3	14.8	0.004
CUT_SPD (Q)	4.0	1	4.0	2.8	0.130
1L by 2L	8.0	1	8.0	5.5	0.043
1L by 2Q	42.7	1	42.7	29.5	0.000
1Q by 2L	2.7	1	2.7	1.8	0.207
1Q by 2Q	8.0	1	8.0	5.5	0.043
Error	13.0	9	1.4		
Total SS	124.0	17			

Partitioned ANOVA

- Tool life study from Montgomery example 6-5
- Note expanded effects
- Plot of two way interactions

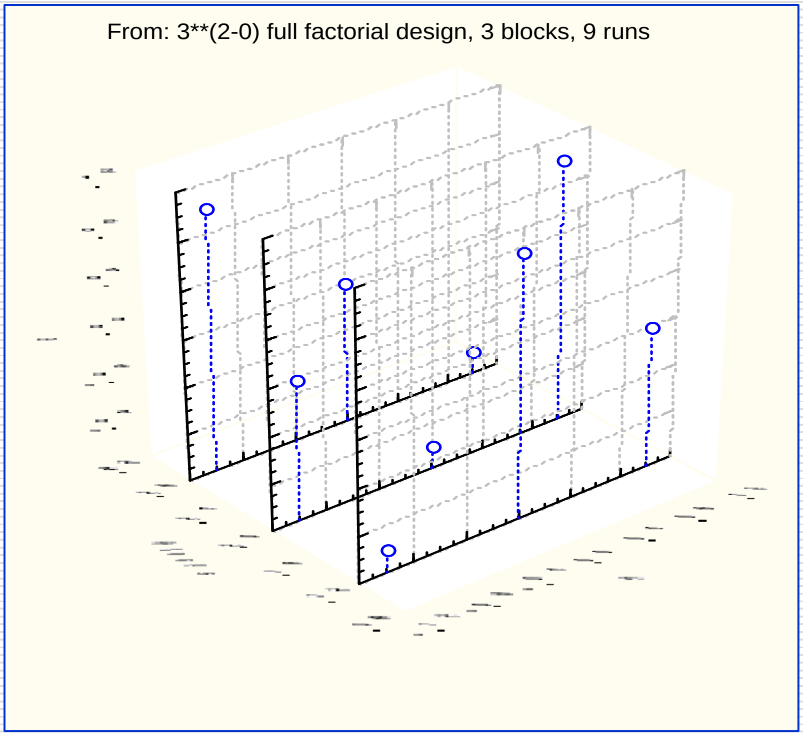


The 3^3 design



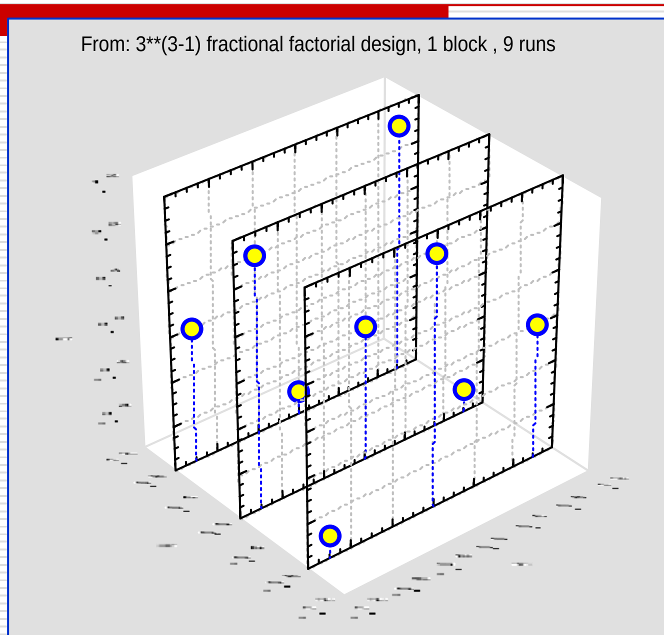
- For 3^3 design -- 27 runs with 26 df
- Main effect sum of squares partitioned as before
- Two ways partitioned as before
- Three ways with (a-1)*(b-1)*(c-1)=8 df -- LxLxL, LxQxL, LxQxQ, QxLxL, etc...
  - (Usually combine three ways and higher into single interaction term since these are difficult to interpret)
- If one or more factors qualitative -- create quadratic models for each of the levels of the qualitative variable

Confounding in 3^k Design



- 3^k so large that confounding in blocks or fractionation often physically required 3^5 is 243 runs (almost 8x equivalent 2^5 design
- 3^k confounded in 3^p ($p<k$) blocks -- 3 blocks, 9 etc
- Confound two df effect with blocks -- e.g. AB^2
- Usual defining contrast method used.
- Can do 3^k factorial in 9 blocks as well

Fractionating 3^k design



One third fraction of 3^3 design

- As in P-B designs, have to deal with partial aliasing of partially quadratic effects.

Correlations of Effects (monty 6-5 and 10-1 3^2.sta)

EXPERIM. DESIGN: 3**(3-1) fractional factorial design, 1 block , 9 r

Factor	C Linea	C Quadra	1L*2L	1L*2Q	1Q*2L	1Q*2Q	1L*3L	1L*3Q
(1)A (L)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A (Q)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(2)B (L)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-.71
B (Q)	0.00	0.00	0.00	0.00	0.00	0.00	-.71	0.00
(3)C (L)	1.00	0.00	0.00	-.71	-.71	0.00	0.00	0.00
C (Q)	0.00	1.00	-.71	0.00	0.00	.71	0.00	0.00
1L*2L	0.00	-.71	1.00	0.00	0.00	0.00	.50	0.00
1L*2Q	-.71	0.00	0.00	1.00	0.00	0.00	0.00	-.50
1Q*2L	-.71	0.00	0.00	0.00	1.00	0.00	0.00	.50
1Q*2Q	0.00	.71	0.00	0.00	0.00	1.00	.50	0.00
1L*3L	0.00	0.00	.50	0.00	0.00	.50	1.00	0.00
1L*3Q	0.00	0.00	0.00	-.50	.50	0.00	0.00	1.00
1Q*3L	0.00	0.00	0.00	.50	-.50	0.00	0.00	0.00
1Q*3Q	0.00	0.00	.50	0.00	0.00	.50	0.00	0.00
2L*3L	0.00	0.00	.50	0.00	0.00	.50	.50	0.00
2L*3Q	0.00	0.00	0.00	.50	-.50	0.00	0.00	-.50
2Q*3L	0.00	0.00	0.00	-.50	.50	0.00	0.00	.50

Alias structure is complex

Running one of the three blocks of a 3^{3-1}

$A=A(AB^2C^2)=A^2B^2C^2=ABC$

$A=A(AB^2C^2)^2=A^3B^4C^4=BC$

$B=B(AB^2C2)=AB^3C^2=AC^2$

$B=A(B^2C^2)^2=A^2B^5C^4=ABC^2$

$C=C(AB^2C^{2)=}AB^2C^3=AB^2$

$C=C(AB^2C^2)=A^2B^4C^5=AB^2C$

$AB=AB(AB^2C^2)=A^2B^3C^2=AC$

$AB=AB(AB^2C^2)^2=A^3B^5C^4=BC^2$

or

$la = A+BC+ABC$

$lb = B+AC^2+ABC^2$

$lc = C+AB^2+AB^2C$

$lab = AB+AC+BC^2$

- Since main effects are aliased with 2 factor -- RIII
- If any two factor interactions are large -- difficult to isolate with these designs
- one-ninth designs more complex.
- No simple augmentation schemes (foldover)
- Monty: *"3^{k-p} designs are solution looking for problem"*

Designs at two and four levels

Design of Experiments

DOE

- Relatively simple -- treat four level as 2x2 level factors

X = 1,2,3,4		
A mapping between X and AxB		
	B = -1	**B = 1**
A = -1	X = 1	X = 3
A = 1	X = 2	X = 4

A full design

Run	A	B	=X	C	D	AB	AC	BC	ABC	AD	BD	ABD	CD	ACD	BCD	ABCD
1	-1	-1	1	-1	-1	1	1	1	-1	1	1	-1	1	-1	-1	1
2	1	-1	2	-1	-1	-1	-1	1	1	-1	1	1	1	1	-1	-1
3	-1	1	3	[illegible]	[illegible]	[illegible]	[illegible]	-1	1	1	-1	1	1	-1	1	-1
4	1	1	4	[illegible]	[illegible]	[illegible]	[illegible]	-1	-1	-1	-1	-1	1	1	1	1
5	-1	-1	1	1	-1	1	-1	-1	1	1	1	-1	-1	1	1	-1
6	1	-1	2	1	-1	-1	1	-1	-1	-1	1	1	-1	-1	1	1
7	-1	1	3	1	-1	-1	-1	1	-1	1	-1	1	-1	1	-1	1
8	1	1	4	1	-1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1
9	-1	-1	1	-1	1	1	1	1	-1	-1	-1	1	-1	1	1	-1
10	1	-1	2	-1	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1
11	-1	1	3	-1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1
12	1	1	4	-1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	-1
13	-1	-1	1	1	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1
14	1	-1	2	1	1	-1	1	-1	-1	1	-1	-1	1	1	-1	-1
15	-1	1	3	1	1	-1	-1	1	-1	-1	1	-1	1	-1	1	-1
16	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1

Sample Analysis

ANOVA Table

Source	**SSQ**	**df**	**MS**
X	SSA+SSB+SSAB	3	MSX
C	SSC	1	MSC
D	SSD	1	MSD
CD	SSCD	1	MSCD
XC	SSAC+SSBC+SSABC	3	MSXC
SSXD	SSAD+SSBD+SSABD	3	MSXD
SSXCD	SSACD+SSBCD+SSABCD	3	MSXCD
Error	SSE	0	MSE
Total	SST	15	

Objective of these last topics

- There are a number of other DOE Topics we have not addressed
- These other topics are occasionally useful
- You should be aware of them and know where to find them
- If needed – two alternatives:
  - Teach yourself (already have foundations)
  - Pete V's Rule 1 of Analysis:

    Seek professional help!

Mainline Classes of Designs

Design of Experiments

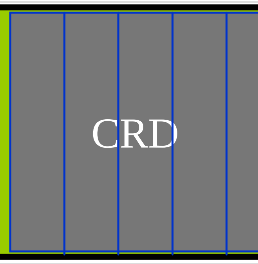
DOE

Factorial Var

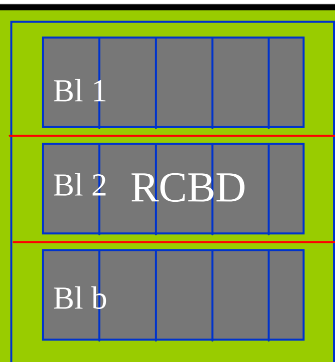
Blocking Var

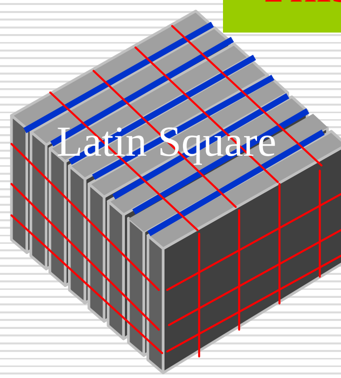
This Course

One Way Designs

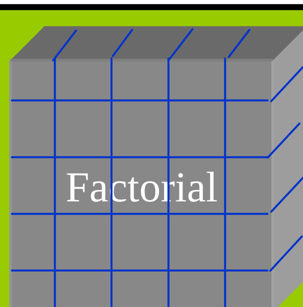


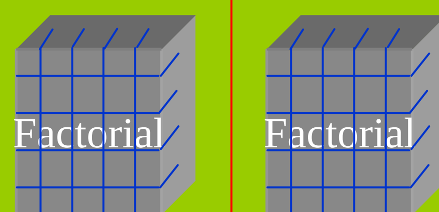
One Way With Blocks





M Way Designs





Factorials in Blocks

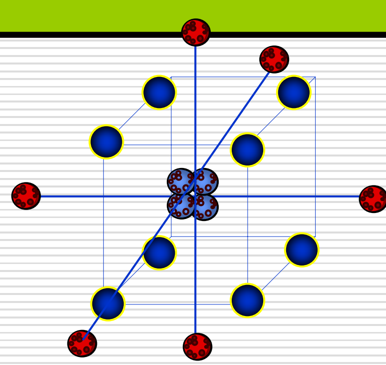
M Way Screening Designs

			High Altitude		Low Altitude	
ECM	Offset	Clutter	Auto Trk	Man Trk	Auto Trk	Man Trk
Wet	Short	High	12.1			5.6
		[illegible]		[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
		[illegible]	[illegible]	[illegible]	[illegible]	3.4
Dry	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
		Low	14.7			4.6
	Long	High	12.2			7.8
		Low		13.6	3.4	

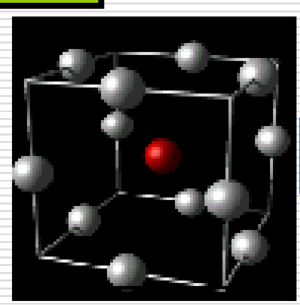
Fractional Factorials

(And Blocks)

M Way Response Surface Designs



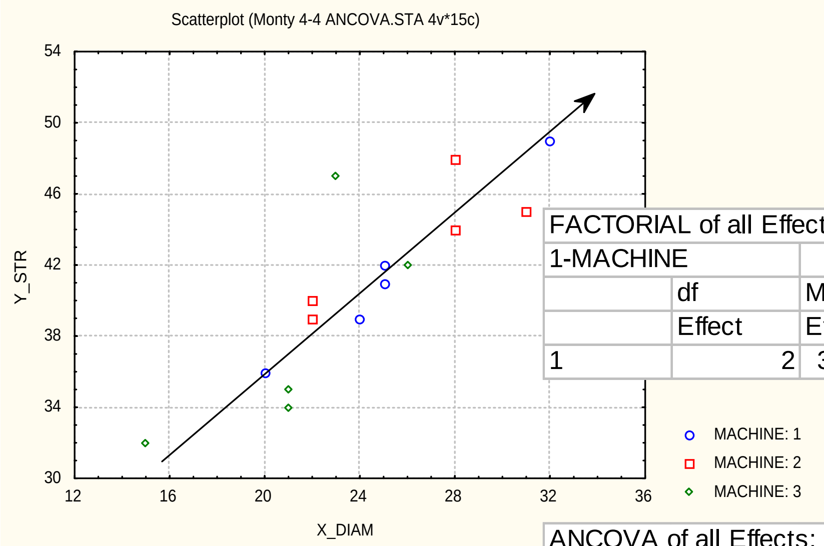
Central Composite



Box-Behnkin

(And Blocks)

ANCOVA -- Analysis of Covariance



FACTORIAL of all Effects; design: (monty 4-4 ancova.sta)						
1-MACHINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	34.86666	12	25.5	1.36732	0.291776

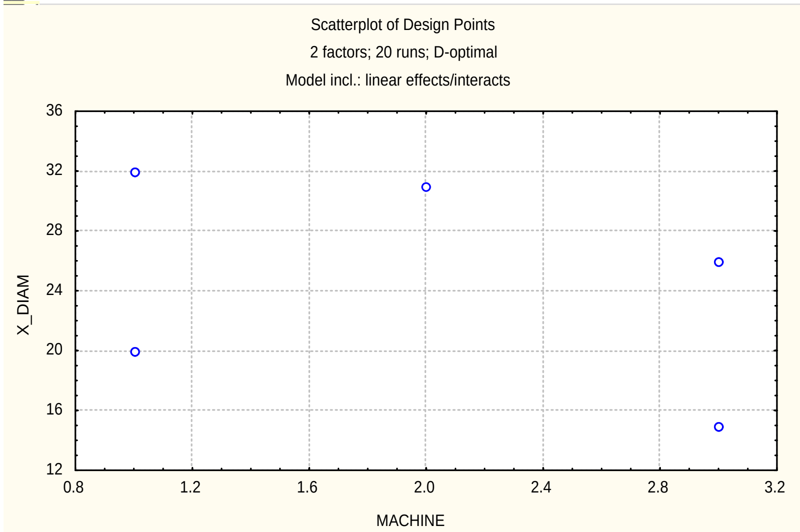
ANCOVA of all Effects; design: (monty 4-4 ancova.sta)						
1-MACHINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	0.874629	11	8.870948	0.098595	0.906901

- Used when some X is random, uncontrollable, but measurable and systematically causes Y to vary with it.
- Examples -- C_d with Mach and MD or TE with Range
- Specify X as a Covariate with Y
- Procedure -- Adjust Y to account for X’s effect on it . Combines ANOVA and Regression to reduce MSE
- With ANCOVA -- attribute cause of variation to appropriate factor -- diameter in Monty 4-4.

Alpha-Optimality

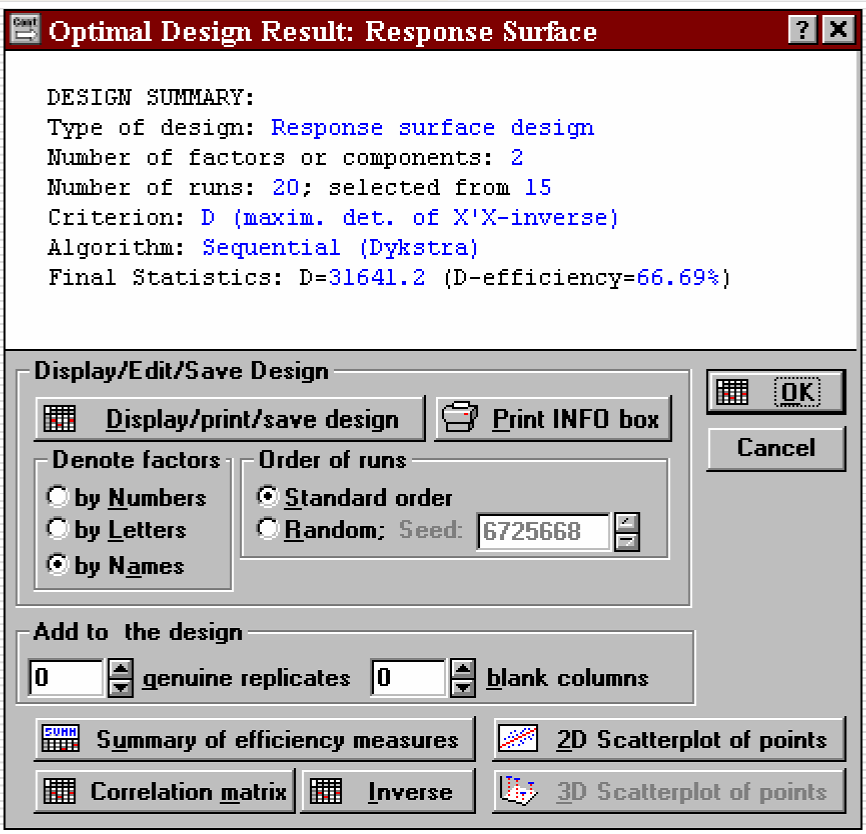
- Factorial designs are (by construction) largely orthogonal
- These are a class of designs that attempt to optimize information from an experiment without full orthogonality
- You specify either: model to fit or points to run
- D Optimal: a D-optimal design will maximize the orthogonality of the the design matrix (the determinant)
- A-Optimal maximizes the diagonal while minimizing the off diagonal elements of the design matrix
- Notes:
  - Statistica has 5 algorithms to search for best design
  - Most useful if you are constrained in experimenting (many infeasible regions or required points)
  - Solutions are not unique or guaranteed to be optimal (search algorithm may get stuck in local max/min)
- May be used to *repair* or augment an incomplete experiment with many missing data points -- which small subset of points to run next.

Example -- D-Optimal Design



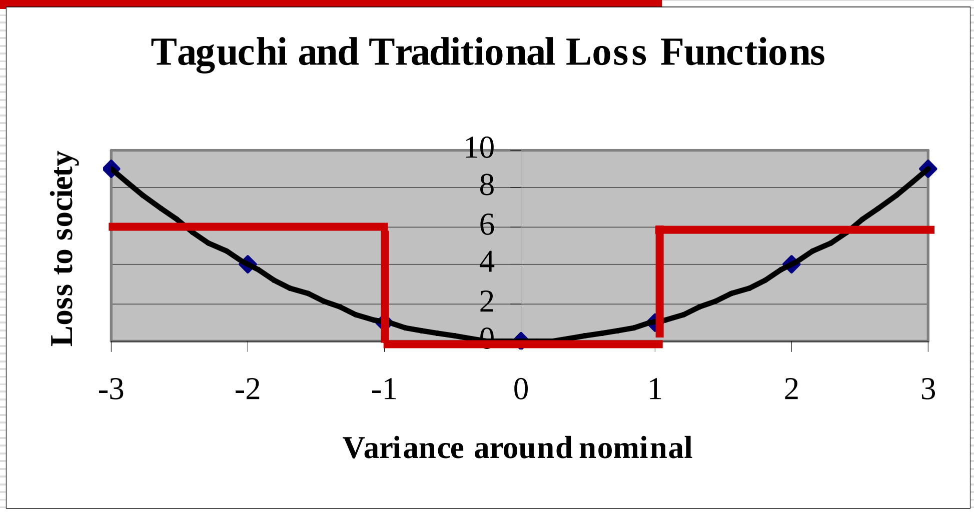
Machine	Diam
1	20
2	30
3	25

Force these points



- Note that all classical designs ***are*** optimal by these criteria

Montgomery's Critique of Taguchi



- Taguchi's quality philosophy:
  - minimum variation around nominal
  - robust to environmental or parameter variance
- Taguchi's Design philosophy
  - interactions are uncommon and unimportant
- Taguchi's analysis philosophy
  - Use S/N ratio's and "pick the winner" marginal means plots
- Summary:
  - Taguchi quality philosophy commendable
  - Taguchi design and analysis practices inefficient, ineffective, and needlessly complicated.

Taguchi L9 Orthogonal Array

Alias Structure (monty 4-4 ancova.sta)

EXPERIM. DESIGN: L9: 4 factors; all factors have 3 levels
* = partially or completely confounded

Effect	1 2	1 3	1 4	2 3	2 4	3 4
1				*	*	*
2		*	*			*
3	*		*		*	
4	*	*		*		

- 2 designs -- Compare alias structures
  - 2^{4-1} R_{IV} half fraction in 8 runs
  - Taguchi L9 4 factors in 9 runs
- Conventional fractional factorials or Plackett-Burman’s have clearer alias structures than Taguchi OA’s and usually a foldover or complementary run approach to continue experimentation

Aliasing of Effects (Computed from Generators)

EXPERIM. DESIGN: 2**(4-1) design
(Factors are denoted by numbers)

Factor	1234
1	234
2	134
3	124
4	123
12	34
13	24
23	14
14	23
24	13
34	12

Nested Designs -- Formulation

Variable A:		Supplier Jones				Supplier Smith		
Variable B:	Sample 1	Sample 2	Sample 3	Sample 4	Sample 1	Sample 2	Sample 3	Sample 4
					Sample 5	Sample 6	Sample 7	Sample 8

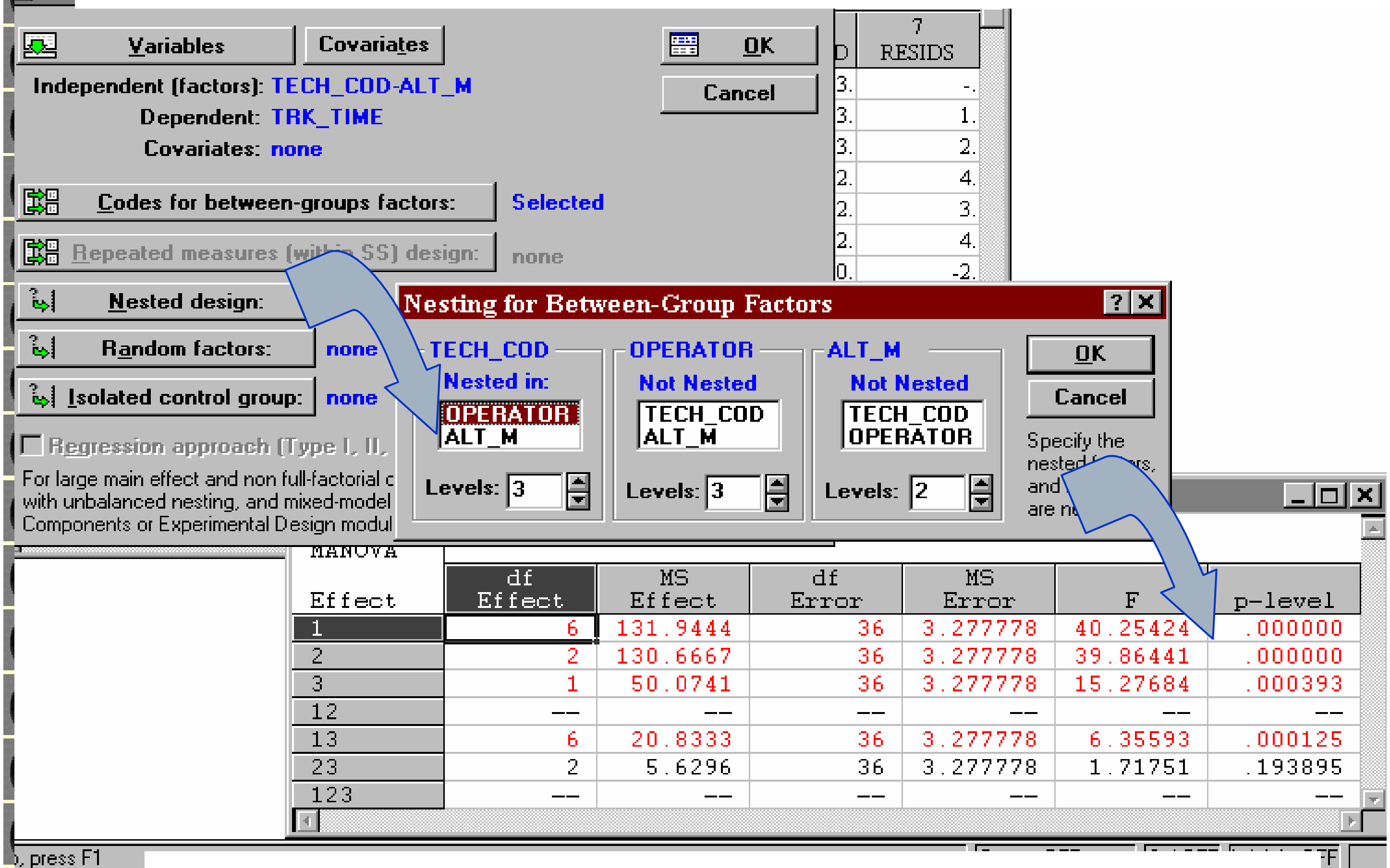
Jones Sample 1 <> Smith Sample 1

- Levels of one factor (e.g. B) are similar but not identical for levels of another factor (e.g. A).
- Montgomery's example -- purity of raw material from two suppliers -- three measurements each.
- Statistica example: test of four English classes -- two from each high school
- Instead of ***crossed*** (as in factorial) B is said to be ***nested*** in variable A
- If factor can be renumbered as in purity example, design is nested.

Variable A:	Park Hills H.S.		Fairborn Baker H.S.	
Variable B:	Ms. Jones	Ms. Pahl	Mr. Frank	Mr. Ernest

Nested Designs -- Analysis

Interpretation: techniques differ in effects by altitude, major differences between operators -- implies select technique by operator and altitude



Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	6	131.9444	36	3.277778	40.25424	.000000
2	2	130.6667	36	3.277778	39.86441	.000000
3	1	50.0741	36	3.277778	15.27684	.000393
12	--	--	--	--	--	--
13	6	20.8333	36	3.277778	6.35593	.000125
23	2	5.6296	36	3.277778	1.71751	.193895
123	--	--	--	--	--	--

- In balanced nested design, there are *a* levels of A, *b* levels of B nested in A
- Suppose you had custom techniques for each operator in an ECM design -- Technique nested in Operator. In ANOVA module, select *All Codes* for each factor then specify *Nesting*
- Note there is no interaction between the nested factors (Operator and Technique) in a nested design

Split Plot Designs -- Formulation

- In some block designs we may not completely randomize run order within blocks -- perhaps a *split-plot design*.
- Consider a ground mount with 3 different power sources for threat. Wish to check 4 ECM techniques for each power source. Want 3 replicates but can only run 12 runs in our hour of testing per day.
- Execution: set power, run four techniques. Change power source. Run four techniques.
- Might consider this a factorial with days as blocks but consider restriction on randomization

Power	Technique	Mission	Response
Commercial	A	1	
	B	1	
	C	1	
	D	1	
Diesel	B	1	
	C	1	
	D	1	
	A	1	
Turbine	D	1	
	A	1	
	C	1	
	B	1	
Commercial	C	2	
	D	2	

Techniques randomized ***within*** *Power source -- Each Mission block divided into three Power* ***whole plots*** *-- Techniques are a* ***split plot*** *treatment within whole plot. Impact of time-varying unknown?*

Split Plot Analysis

Blocked Factorial ANOVA Table
A-TECH_COD (3), B-POWER_S (3), C-MSN_BLK (2)

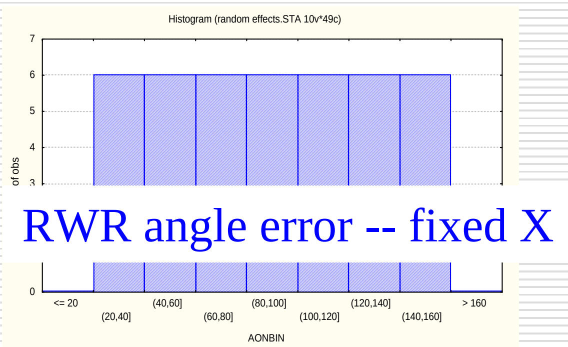
Source	Name	SS Effect	df Effect	MS Effect	F	p-level
A	Tech	100.3	2	50.2	7.6	0.0143
B	Power	76.0	2	38.0	5.7	0.0285
AB		138.7	4	34.7	5.2	0.0228
C	Blocks	18.0	1	18.0	--	
Error	Blk x Tr Interaction	53.0	8	6.625		
AC		20.3	2	10.2		
BC		12.0	2	6.0		
ABC		20.7	4	5.2		
Total		386.0	17			

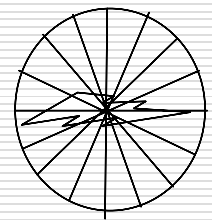
Split Plot ANOVA Table
A-TECH_COD (3), B-POWER_S (3), C-MSN_BLK (2)

Source	Name	SS Effect	df Effect	MS Effect	Split Plot F	Factorial F	Split Plot p
B	Main TR Power	76.0	2	38.0	6.3	5.7	0.1364
C	Msn Blocks	18.0	1	18.0	--		
BC	Main Error	12.0	2	6.0	--		
A	Subplot TR - Tech	100.3	2	50.2	9.7	7.6	0.0292
AB	Main x Sub (PxT)	138.7	4	34.7	6.7	5.2	0.0461
AC	Sub x Blk	20.3	2	10.2	--		
ABC	Subplot Error	20.7	4	5.2	--		
Total		386.0	17				

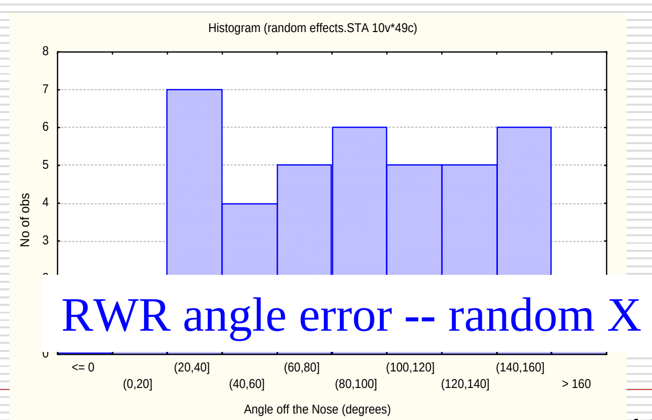
- In blocked factorial -- blocks interactions estimate error
- In Split-plot, two error estimates -- among whole plots and within subplots
- keep this model in mind when factors are difficult to change and you restrict randomization within blocks. May cause you to re-analyze data as split plot.

Fixed vs. Random Effects Model Formulation





- In fixed effects model, only Y is a random variable -- we choose and *fix* X at *a* levels
- In random effects model, we *randomly* choose *a* levels of X from a large number of values
- In *random effects model*, both Y and X are random
- In *mixed* model, some X are fixed, some random
- Example -- actual angle when df error measured in RWR tests. Angle is a random factor rather than same angle each pass



Random Effects Statistical Model

Design of Experiments DOE

Model: $SS_T = SS_{Treatments} + SS_{Error}$

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Fixed Effects Model

$$V(y_{ij}) = \sigma^2$$

$$H_O : \tau_i = 0$$
$$H_a : \tau_i \neq 0$$

$$E(MS_{Treatments}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau^2{}_i}{a-1}$$

Random Effects Model

$$V(y_{ij}) = \sigma_\tau{}^2 + \sigma^2$$

$$H_O : \sigma_\tau{}^2 = 0$$
$$H_a : \sigma_\tau{}^2 > 0$$

$$E(MS_{Treatments}) = \sigma^2 + n\sigma_\tau{}^2$$

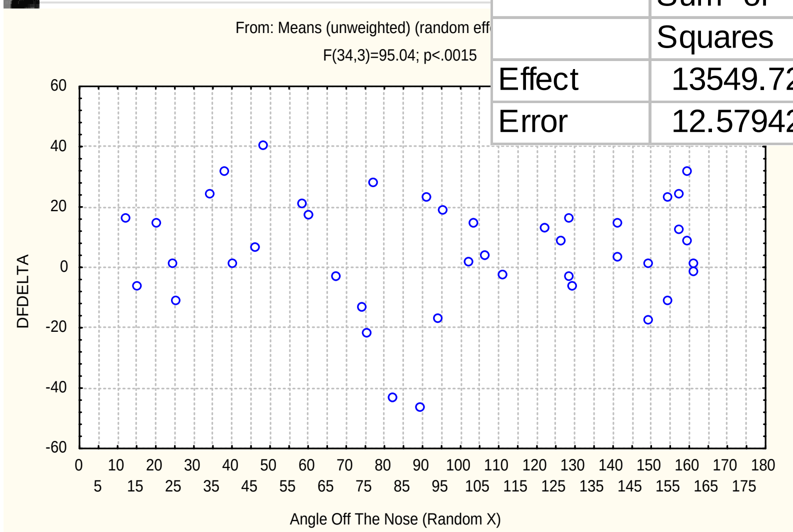
$$E(MSE) = \sigma^2$$

$$F_o = \frac{MS_{Treatments}}{MS_{Error}}$$

- Random effects also known as *Components of Variance* model -- part to X, part to Error
- Inferences concerning one level of X meaningless;
- Inferences apply to entire population of X

Random Effects Model -- Analysis

MAIN EFFECT: AONTRUNC (random effects.sta)					
Error: Within					
	Sum of		Mean		
	Squares	df	Square	F	p-level
Effect	13549.72	34.00	398.5213	95.04	0.0015
Error	12.57942	3.00	4.193141		



$$SST = SS_{treatments} + SS_E$$

$$H_o: \ \sigma_\tau = 0$$

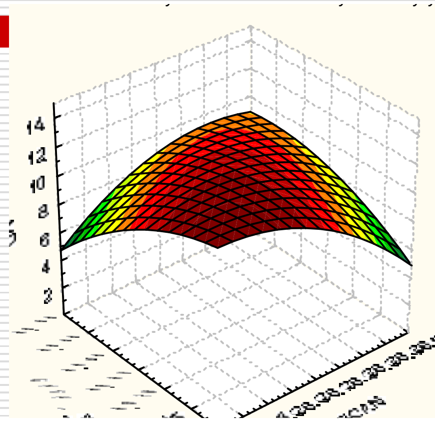
$$H_a: \ \sigma_\tau > 0$$

- We must account for both variance of X and the variance of Y in our model
- No interaction between predictors when X is random or mixed
- No estimate of individual cell means for levels of X
- Simply -- are some levels of X different than others?

Analyzed as a fixed effects model

Summary of all						
1-AONCODE, 2-SIDECODE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	6	623.21	28	55.82	11.16	0.000002
2	1	7311.94	28	55.82	130.99	0.000000
12	6	420.69	28	55.82	7.54	0.000071

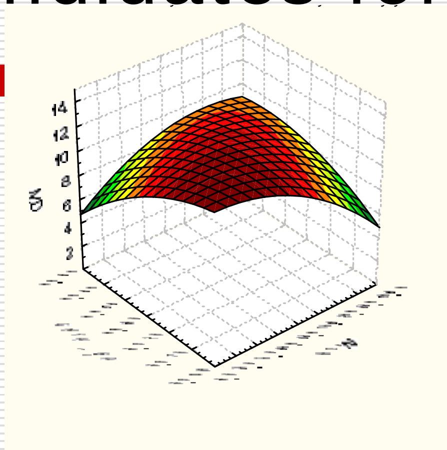
Intro to Response Surfaces

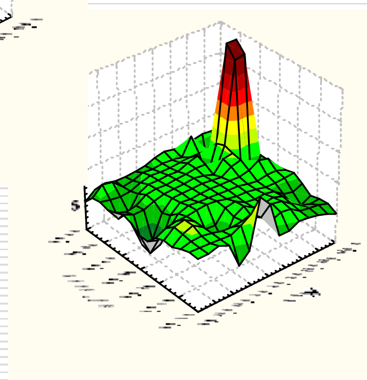


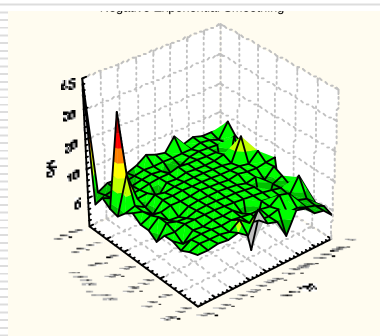
- Response Surface Methodology -- a collection of math/stat tools to optimize settings of quantitative predictor (X) variables on Y

  $Y = f(X_1, X_2, \ldots X_k) + \varepsilon$
- Example -- effect of temperature and pressure on chemical process or effect of scan rate offset and duty cycle on track errors
- General procedure is
  - start at a region
  - estimate first order
  - climb
  - center second matrix
    - If curvature, discover min/max points
    - If not -- climb again and iterate

Candidates for RSM







- Quantitative predictors and response variables
- some indication of gradients to exploit
- When an empirical model is desired
- Start with as many variables as desired
- Cull to less than 6 predictor variables (interactions are harder) -- best if 3 or 4
- In my Wing
  - Modeling and Simulation
  - HWIL (EW Ground mounts)
  - AFEWES and other installed test fac’y
  - Sys Integration Lab testing

Typical RSM Models

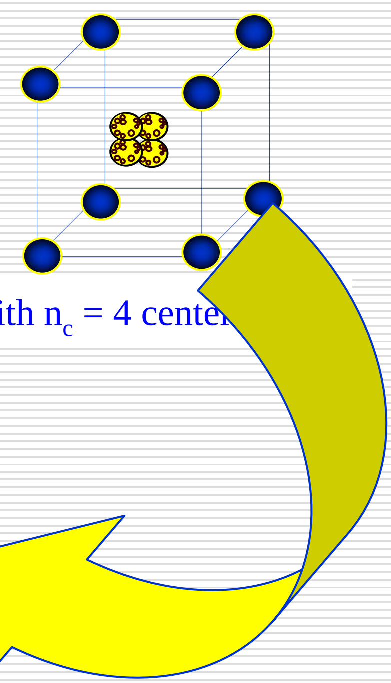
First Order Model

$$y = \beta o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

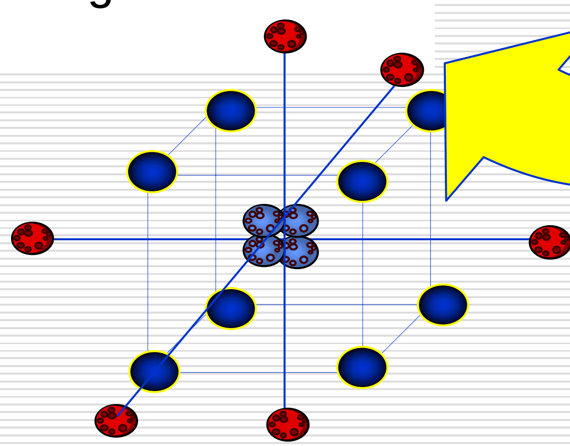
Second Order Model

$$y = \beta o + \beta_1 x_1 + \beta_{11} x^2{}_1 + \beta_2 x_2 + \beta_{22} x^2{}_2 + \beta_{ij} x_i x_j + \varepsilon$$

- First order model for first check of region (with center points)
- Use center points to estimate
  - interactions,
  - errors, and
  - curvature
- If needed, fit RSM model (CCD or B-B) to local region

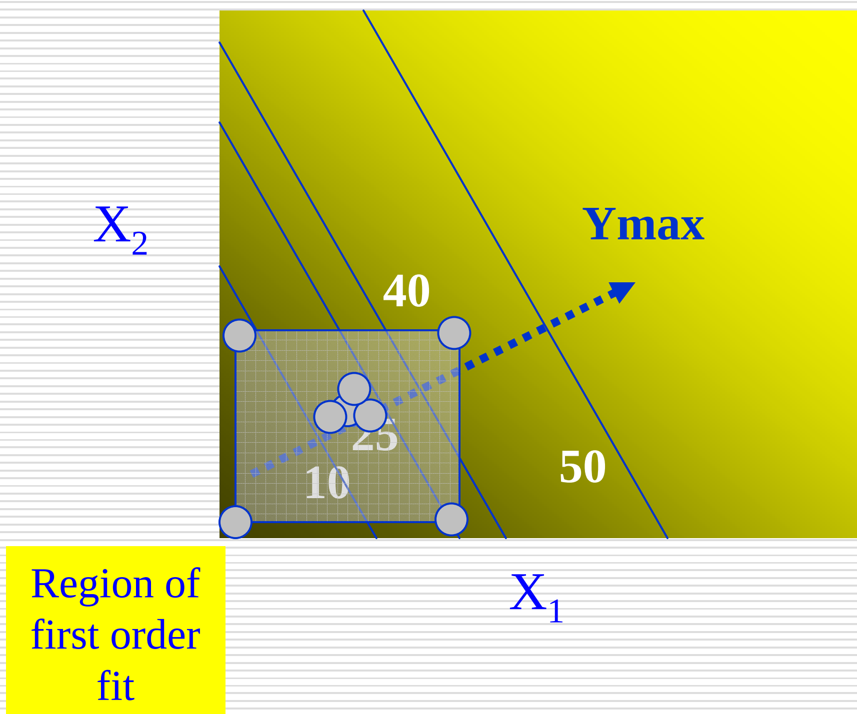


2^3 with n_c = 4 center



Central Composite

Method of Steepest Ascent



- Steps along the path are proportional to the first derivative of the first order model (slopes of the variables)
- Method usually uses coded variables and transforms back and forth

Transform Coded

$$x_{coded} = \frac{x_n - \bar{x}_n}{x_{\max} - x_{\min}}$$

Step size $x_i = \Delta x_i$ from Max$\left|\hat{\beta}_j\right|$

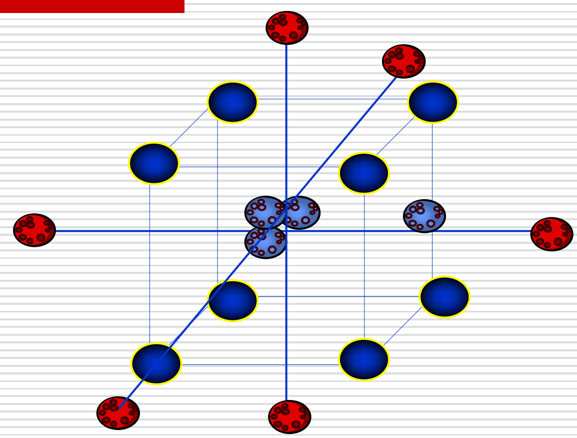
Steepest Ascent Units

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_i}$$

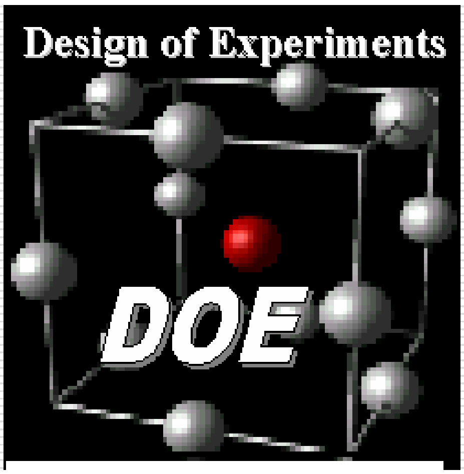
Second Order Design Properties

- We Want
  - reasonable distribution of data points for robustness and scope
  - allows model adequacy including curvature and lack of fit to be investigated
  - allows experiments in blocks
  - allows higher-order designs to be built up sequentially
  - provides an internal estimate of error
  - requires as few points as possible
  - enough, but not too many levels of the independent variable
  - ensures simplicity of calculations
  - is orthogonal
  - is rotatable

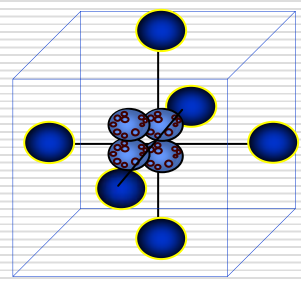
Second Order Orthogonal Designs



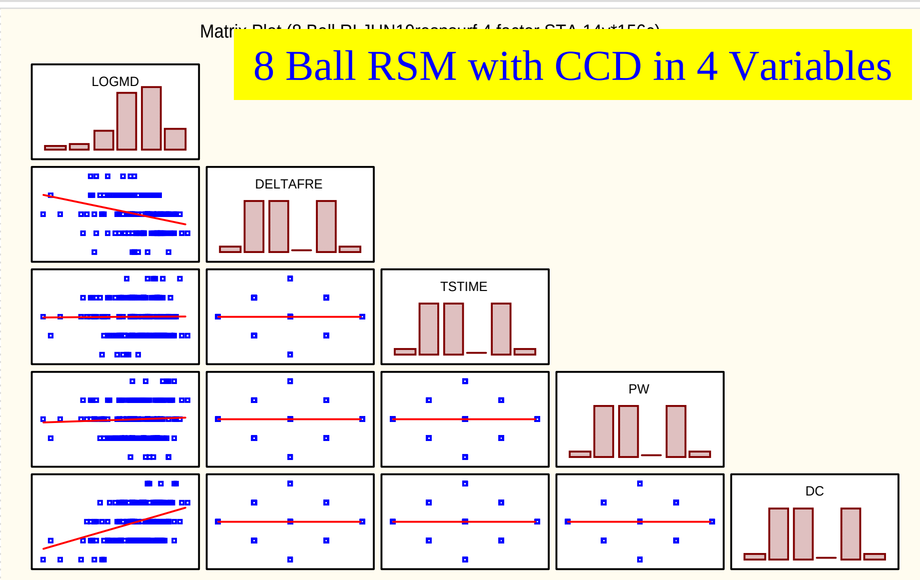
Central Composite



Box Behnken Spherical Edge-Centered Design



Face-Centered CCD



8 Ball RSM with CCD in 4 Variables

The *Least* You Should Know…

- 3-level designs seem attractive when curvature is desired
- Drawbacks include:
  - larger designs (3^k)
  - difficult to interpret quadratic interactions
  - undesirable fractional confounding patterns
- **Summary** -- 3^k designs are a "*solution looking for a problem*"

- Mixed 2 and 3 level designs can be represented in 2k model
  - may need to roll own mixed design and let Statistica analyze it
- Mixed models with four level variables can be represented with 2x2-level variables

The *Least* You Should Know…

- Variations on factorial designs and ANOVA analysis are many
- You know the mainline approaches and have the math to learn others
- If B_i differs with A settings, you have a ***nested*** design
- ***Split plot*** designs analyze restricted randomization within blocks
- If X is chosen (or observed) randomly, the model is a ***random effects*** model
- When X is uncontrollable, affects Y, and can be observed, it may be a ***covariate*** analyzed by ANCOVA
- ***Alpha-optimality*** (D-, A-, G-) assumes substantial process knowledge. May be used for repair or constrained situations
- ***Taguchi*** made strong contributions to quality improvement; there are better DOE approaches than Taguchi's, however.

The *Least* You Should Know…

- RSM is a straight-forward extension of factorial designs -- with a navigation tool
- Good RSM candidates have 3-5 continuous variables with a gradient to exploit
- Basic RSM method:
  - 1st order factorials with cp to estimate gradient
  - ascend gradient to top (15-20 runs)
  - run factorial with cp's to detect curvature near optimum
  - augment an RSM design to estimate model
  - Confirm optimum with factorial near optimal value
- Consider either central composite design or Box-Behnkin design for second order models